Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions

By Alvin E. Roth[1]
This draft: 13 June, 2007

Abstract:

The deferred acceptance algorithm proposed by Gale and Shapley (1962) has had a profound influence on market design, both directly, by being adapted into practical matching mechanisms, and, indirectly, by raising new theoretical questions. Deferred acceptance algorithms are at the basis of a number of labor market clearinghouses around the world, and have recently been implemented in school choice systems in Boston and New York City.

In addition, the study of markets that have failed in ways that can be fixed with centralized mechanisms has led to a deeper understanding of some of the tasks a marketplace needs to accomplish to perform well. In particular, marketplaces work well when they provide thickness to the market, help it deal with the congestion that thickness can bring, and make it safe for participants to act effectively on their preferences. Centralized clearinghouses organized around the deferred acceptance algorithm can have these properties, and this has sometimes allowed failed markets to be reorganized.

[1] Prepared for Gale's Feast: a Day in honor of the 85th birthday of David Gale, July 2007, Stony Brook. Acknowledgments: It is my very good fortune to have been able to incorporate the insights of Gale and Shapley (1962) in so much of my own work. In this I have been exceptionally lucky in my colleagues and coauthors. For help specifically in preparing this paper I particularly thank my old friend and collaborator Elliott Peranson for new details on the history of the medical residency match. This paper has also benefited from Fuhito Kojima's careful reading and wide acquaintance with the literature.

Introduction:

Matching is one of the important functions of markets. Who gets which jobs, which school places, who marries whom, these help shape lives and careers.

Partly for this reason, a substantial literature has grown out of the remarkable paper "College Admissions and the Stability of Marriage" that David Gale and Lloyd Shapley published in *The American Mathematical Monthly* in 1962 (henceforth GS). In that short and almost non-technical paper they proposed a simple model of two-sided matching, in which men and women (or students and colleges) each had preferences over individuals to whom they might be matched in the other set. They proposed an algorithm for finding a "stable" matching that works by having agents on one side of the market make proposals (offers or applications) to agents on the other, in order of preference. Those who receive more proposals than they can accept reject their least preferred, but don't immediately accept those they don't reject; they instead hold them without commitment, and acceptances are deferred until the end of the algorithm. In the meantime, agents who have been rejected make new proposals, which lead to new rejections (including of proposals that were held at an earlier period but are less preferred than a new proposal), until there are no rejected agents who wish to make further proposals. At this point all proposals that are being held are finally accepted, to produce a matching.

GS used this deferred acceptance algorithm to prove not only that the set of stable outcomes is nonempty for all preferences of the kind they considered, but that, when preferences are strict, there always exists, for each side of the market, a stable matching that is optimal for agents on that side of the market in the surprisingly strong sense that every agent likes the optimal stable matching for his/her side of the market as well as any stable matching. They showed in passing that the non-emptiness of the core is related to the two-sidedness of the market. And, after appropriate caveats about the abstraction of their model from actual college admissions, they stated in closing that (p14) "It is our opinion, however, that some of the ideas introduced here might usefully be applied to certain phases of the admissions problem."

This opinion has turned out to be well justified. In the last few years, admissions to high schools in New York City, and to all public schools in Boston, have been reorganized into clearinghouses that use deferred acceptance algorithms adapted to local needs, which solve serious problems encountered by the previously used systems. And it turns out that well before GS, similar ideas had been incorporated, in the early 1950's, into the successful clearinghouse through which most American doctors still obtain their first jobs. We now know that deferred acceptance algorithms have been independently discovered a number of times, in part because they capture a "folk idea" of how markets operate. In turn, seeing how clearinghouses operate differently from the less centralized marketplaces they replace has yielded new insights into the tasks that markets have to perform to work well, whether they are centralized or decentralized. And the design of clearinghouses, and of rules for decentralized markets, has also raised new theoretical questions.

This review will be organized as follows. Section 1 reviews the results presented in GS. Section 2 selectively reviews some subsequent theoretical developments that helped set the stage for making these results practical tools for analyzing and designing market mechanisms. Section 3, again selectively, briefly reviews some of the history of applications of deferred acceptance algorithms in markets, including some new historical details regarding the medical resident match, that came to light in recent discussions. In addition to medical labor markets, I'll discuss matching children to public schools. Labor markets and school choice are the applications that have led most directly to the new theoretical developments and open questions that I will discuss in Section 4. Section 5 will discuss how these developments relate to insights into the detailed working of markets, both centralized and decentralized, that have grown from the study of market failures and how they can be fixed. An Appendix fills in some intellectual history.

The literature that derives from GS has grown a great deal since Marilda Sotomayor and I attempted a comprehensive survey in Roth and Sotomayor (1990). The Science Citation Index recently showed 434 citations to GS, 343 of them published after 1990.[2] The present review does not attempt even a cursory survey of the whole landscape of new developments, but rather simply tries to trace a path from the original theory, through some empirical and practical aspects of subsequent studies and designs of market clearinghouses, to some new theoretical questions and research directions.

1. The marriage and college-admissions models

GS presented two closely related models of two-sided matching, differing only in whether the agents on each side each wish to be matched with *one* other agent of the opposite sort (the "marriage model"), or whether the matching is many-to-one (the "college admissions" model).

In the marriage model there are two disjoint sets of agents, $M = \{m_1,..., m_n\}$ and $W = \{w_1,..., w_p\}$, Men and Women, each of whom has complete and transitive preferences over the individuals on the other side (and the possibility of being unmatched, which we model as being "matched to yourself"). Preferences can be represented as rank order lists of the form $P(m_i) = w_3, w_2, ... m_i$, denoting that man m_i's first choice is w_3, his second choice w_2 [$w_3 >_{mi} w_2$] and so on, until at some point he prefers to remain unmatched (i.e. matched to himself). If agent k (on either side of the market) prefers to remain single rather than be matched to agent j, i.e. if $k >_k j$, then j is said to be *unacceptable* to k, and in examples preferences will be represented just by the list of acceptable mates. If an agent is not indifferent between any two acceptable mates, or between being matched to an acceptable mate and being unmatched, his or her preferences will be called *strict*.

[2]Further play with the citation index reveals that five authors have cited Gale and Shapley in ten or more papers: Roth, Sönmez, Sotomayor, Irving, and Balinski, in 78 papers in various combinations and with 35 other coauthors.

An *outcome* of the game is a *matching*: $\mu: M \cup W \rightarrow M \cup W$ such that $w = \mu(m)$ if and only if $\mu(w)=m$, and for all m and w either $\mu(w)$ is in M or $\mu(w) = w$, and either $\mu(m)$ is in W or $\mu(m) = m$. That is, an outcome matches agents on one side to agents on the other side, or to themselves, and if w is matched to m, then m is matched to w.

A matching μ is *blocked by an individual* k if k prefers being single to being matched with $\mu(k)$, i.e. $k >_k \mu(k)$. A matching μ is *blocked by a pair of agents* (m,w) if they each prefer each other to the partner they receive at μ, i.e. $w >_m \mu(m)$ and $m >_w \mu(w)$.

A matching μ is *stable* if it isn't blocked by any individual or pair of agents.

In the marriage model a stable matching is efficient, and the set of (pairwise) stable matchings equals the core of the game whose rules are that agents from opposite sides of the market can match if and only if they both agree.

Theorem 1 (GS): A stable matching exists for every marriage market.

GS proved the theorem by presenting and analyzing the following algorithm.

GS Deferred Acceptance Algorithm, with men proposing

0. *If some preferences are not strict*, arbitrarily break ties (e.g. if some m is indifferent between w_i and w_j, order them consecutively in alphabetical order. Different agents may break ties differently: e.g.. tie-breaking can be random, or decentralized by having each agent fill out a strict preference list.)

1 a. Each man m proposes to his 1st choice (if he has any acceptable choices).
b. Each woman rejects any unacceptable proposals and, if more than one acceptable proposal is received, "holds" the most preferred and rejects all others.

k a. Any man rejected at step k-1 makes a new proposal to its most preferred acceptable mate who hasn't yet rejected him. (If no acceptable choices remain, he makes no proposal.)
b. Each woman holds her most preferred acceptable offer to date, and rejects the rest.

STOP: when no further proposals are made, and match each woman to the man (if any) whose proposal she is holding.

Note that the proof of the theorem now follows from the observation that the algorithm always stops (since no man proposes twice to the same woman) and the matching produced in this way is itself stable. No man ever proposes to an unacceptable woman, and no woman ever holds the offer of an unacceptable man, and if some man would prefer to be matched to a woman other than his assigned mate, he must, according to the

algorithm, have already proposed to her, and she has rejected him, meaning she has a man she strictly prefers, hence they cannot form a blocking pair.[3]

GS observed that which side of the market proposes in a deferred acceptance algorithm has consequences.

Theorem 2 (GS): When all men and women have strict preferences, there always exists an M-optimal stable matching (that every man likes at least as well as any other stable matching), and a W-optimal stable matching. Furthermore, the matching μ_M produced by the deferred acceptance algorithm with men proposing is the M-optimal stable matching. The W-optimal stable matching is the matching μ_W produced by the algorithm when the women propose.

GS proved Theorem 2 by showing that, in the course of the deferred acceptance algorithm, no man is ever rejected by a woman who could be matched to him at any stable matching.

They also showed that the two theorems depend on the two-sided nature of the problem, by considering, in contrast, a one-sided "roommates" problem in which every agent has preferences over all other agents, and any two agents can be matched together. They considered an example in which there are four potential roommates {1,...4}, such that 1's first choice is 2, 2's first choice is 3, and 3's first choice is 1, and 1-3 all rank 4 as their last choice. Then (regardless of poor 4's preferences), since each of 1,2, and 3 is the first choice of one of the others, no partition of the four potential roommates into two rooms is stable, since whoever is matched with 4 will form a blocking pair with the roommate whose first choice he is.[4]

GS also formulated a many-to-one "college admissions" model populated by a set of colleges and a set of students that was identical to the marriage model except that each college Ci wishes to be matched to qi $\geq$ 1 students, while each student is interested in being matched to only 1 college. That is, each student had a preference over colleges, each college had a preference over individual students, and a matching was a function that assigned each student to no more than one college, and each college i to no more than qi students. They observed that essentially the same deferred acceptance algorithm

[3] Sotomayor (1996) provides an equally simple proof of the nonemptiness of the set of stable matchings that connects stability with a broader notion of stability with respect to unmatched agents.

[4] While I will not discuss it here, a small literature has grown around the issues of finding conditions in which the set of stable matchings is non-empty for the roommates problem, and the performance of algorithms that can produce them when they exist. There were also, over the years, a number of results that supported the conjecture that Gale and Shapley's main results could not be generalized beyond a two-sided structure. (One of the most elegant of these is the proof by Abeledo and Isaak 1991 that if you model a set of agents as the nodes on a graph, and connect a pair of agents only if they are eligible to be matched to one another, and if a stable matching exists for any preferences, then the graph must be bipartite, i.e. it must be possible to divide the set of agents into two disjoint sets such that agents in one set may only be matched to agents in the other.) However, the surprising generalization of Ostrovsky (2007), which looks at "supply chains" of agents, that are stable if there are no "blocking chains," removes the two-sided restriction in a natural way.

(with college i proposing at each point to its qi most preferred students who hadn't yet rejected it in the college-proposing version, or rejecting all but the qi most preferred applications it had received at any point of the student-proposing version) would produce a stable matching defined as before. That is, the outcome produced by the algorithm would not admit any student-college blocking pairs defined precisely as for the marriage model.

2. Subsequent developments:

GS implicitly raised a number of interesting questions, and other new questions arose as game theory developed and as the model was adapted and used to explore actual markets and design new ones.

2.1 The marriage model

One line of investigation concerns why there are optimal stable matchings for each side of the market. One fruitful line of answers to this question, which generalizes well beyond the simple marriage model, is attributed by Knuth (1976) to John Conway. The essential idea is the following. Suppose everyone has strict preferences and let $\mu \neq \nu$ be two stable matchings. Then if we ask each man m to point to whichever mate he prefers from the two, i.e. to $\mu(m)$ or $\nu(m)$, then no two men point to the same woman (since a woman who was pointed to by two different men would prefer one of them, and would form a blocking pair to one of the matchings with that man, which contradicts the assumption that they are stable matchings). So there is a new matching that every man likes at least as well as both μ and ν. Furthermore, this new matching is itself stable. More generally, let the partial order $>_M$ on the set of stable matchings be defined for matchings $\mu \neq \nu$ by $\mu >_M \nu$ if, for every man m in M, either $\mu(m) >_m \nu(m)$ or $\mu(m) = \nu(m)$. That is, $\mu >_M \nu$ if every man agrees that μ is at least as good as ν.

Theorem 3 (Lattice Theorem (Conway in Knuth, 1976)[5]: When preferences are strict, the set of stable matchings is a lattice with respect to the partial order $>_M$. The maximum element of the lattice is μ_M, the stable matching produced by the men-proposing deferred acceptance algorithm, and its minimum element is μ_W, the matching produced by the women-proposing deferred acceptance algorithm.

The Lattice theorem gives insight into other properties of the set of stable matchings, such as the following result, also generalizable to more complex models. (Compare the set of men who must be matched at the man-optimal stable matching to those who are matched at the woman-optimal stable matching, which by the above theorem is the worst stable matching for every man, and note that the number of matched men and women must be the same at every stable matching…)

[5] For extensions of the lattice structure to more general matching models, see e.g. Blair (1988), Martinez, Masso, Neme and Oviedo (2001), Sotomayor (2000).

Theorem 4 (McVitie and Wilson 1970): the set of unmatched men and women is the same at every stable matching.

Another way to examine the optimal stable matchings is to compare them with unstable matchings:

Theorem 5 (Roth, 1982a): The M-optimal stable matching μ_M is *weakly* Pareto optimal for the men in the set of all matchings (and symmetrically, the W-optimal stable matching for the women). That is, there can be no matching (even an unstable matching) that all men strictly prefer to the man optimal stable matching.

The following example (also from Roth 1982a) shows that the theorem can't be strengthened to strong Pareto optimality, i.e. it may be possible to make some men better off than at μ_M without making other men worse off. This example will also be useful when we discuss the ways in which various results generalize beyond the marriage model, and when we consider the issues that non-strict preferences raise in applications to school choice.

Example 1: Let M = {m1, m2, m3} and W = (w1, w2, and w3} with preferences given by :

P(m1) = w2, w1, w3	P(w1) = m1, m2, m3
P(m2) = w1, w2, w3	P(w2) = m3, m1, m2
P(m3) = w1, w2, w3	P(w3) = m1, m2, m3

Then at μ_M **= [(m1,w1); (m2,w3); (m3,w2)]**, m1 and m3 each receive their second choice mate (while m2 receives his last choice). But at the (unstable) matching μ **= [(m1,w2); (m2,w3); (m3,w1)]**, m1 and m3 each receive their first choice, so they are both strictly better off than at the Man-optimal stable matching μ_M, and m2 is not worse off.

Strategic properties:

A different kind of question, which has proved to be of critical importance in market design, concerns the strategic properties of a market organized via a clearinghouse. If a clearinghouse is organized along the lines of the deferred acceptance algorithm, we can ask whether participants will find it in their interest to provide the private information about their preferences on which the algorithm depends. Note that the algorithm was described above as if people take actions in the course of the algorithm, and we can ask whether those actions would best serve their interests. To put the question another way, is it possible to design a clearinghouse in which a stable matching is produced from participants' stated rank order lists in such a way that it will never be in someone's interest to submit a rank order list different from their true preferences? The following theorems show first that there are limits on what any mechanism that produces stable matchings can accomplish in this regard, and then shows a sense in which the deferred acceptance algorithm performs up to those limits. (A "stable matching mechanism" is a

function from any stated preferences to a matching that is stable with respect to those preferences.)

Theorem 6 Impossibility Theorem (Roth, 1982a):
No stable matching mechanism exists for which stating the true preferences is a dominant strategy for every agent.

Since a stable matching mechanism has to produce a stable matching for every instance of the marriage model, a proof of the Impossibility Theorem can be based on the simplest of examples. Consider the case of two men and two women, with preferences P given by

P(m1) = w1, w2 P(w1) = m2, m1
P(m2) = w2, w1 P(w2) = m1, m2

Then there are only two stable matchings, μ_M = [(m1,w1); (m2,w2)], and μ_W = [(m1,w2); (m2,w1)], and so any stable matching mechanism h must produce one of them when the stated preferences are P, i.e. h(P) equals μ_M or μ_W. Suppose h(P) = μ_M (the argument is symmetric if h(P) = μ_W). Then, if w1 say, were to state that m1 was unacceptable, by submitting the preference P'(w1) = m2 (and if everyone else submitted the preferences P, so that the stated preferences are P' = [P(m1), P(m2), P'(w1), P(w2)]), the resulting matching would have to be h(P') = μ_W since μ_W is the unique stable matching when preferences are P'. So w1 would have profited from this mis-statement of her preferences.

Indeed, anyone who isn't matched to their mate at his or her optimal stable matching can potentially manipulate a stable mechanism in this way.

Theorem 7 (Roth and Sotomayor, 1990): When any stable mechanism is applied to a marriage market in which preferences are strict and there is more than one stable matching, then at least one agent can profitably misrepresent his or her preferences, assuming the others tell the truth. (This agent can misrepresent in such a way as to be matched to his or her most preferred achievable mate under the true preferences at every stable matching under the mis-stated preferences.)

Indeed, any agent who receives a mate less preferred than his or her optimal stable matching could profitably manipulate by truncating his or her preference right after that mate. However it is possible to design the mechanism so that one side of the market can never do any better than to state their true preferences.

Theorem 8 (Dominant strategy theorem: Dubins and Freedman 1981, Roth 1982)
In the game induced by the man-proposing deferred acceptance algorithm, in which each player states a preference list, it is a dominant strategy for each man to state his true preferences.

The original proofs were difficult, but as is not so seldom the case, easier proofs result from the following more powerful theorem.

Theorem 9 (Limits on successful manipulation.) (Demange, Gale, and Sotomayor, 1987). Let P be the true preferences (not necessarily strict) of the agents, and let P' differ from P in that some coalition C of men and women mis-state their preferences. Then there is no matching μ, stable for P', which is preferred to *every* stable matching under the true preferences P by all members of C.

Dubins and Freedman in fact had proved something between Theorems 8 and 9, namely that no coalition of men could collectively manipulate in such a way as to strictly improve all of their mates in comparison to μ_M.

However, when we consider which of these results are robust to the generalizations that will be necessary for models that can be applied to real economic environments, we find that, while the possibility of dominant strategy mechanisms for individuals remains, the hope that such mechanisms can also prevent coalitions from profitably colluding by mis-stating their preferences vanishes.

To formalize this we will have to consider more general models, with money and/or many-to-one matching, but a good idea of what will generalize and what will not can be obtained by reexamining Example 1. The reason that m1 and m3 don't get their first choices at μ_M in Example 1 is because of competition with m2. If m2 were to mis-state his preferences so that only w3 was acceptable to him, i.e. if he were to state his preferences as P'(m2) = w3, then, if everyone else stated their preferences as before, the outcome μ would be the result of the deferred acceptance algorithm. As Theorem 4 implies, this can't help m2, but in this case it wouldn't hurt him either, as he was matched to w3 even at μ_M, i.e. even when he stated his true preferences. But his mis-statement helps m1 and m3, who now get their first choices. So, in a model in which there was the possibility of transferring any money at all between agents, m1 or m3 could give a penny to m2 from their gains, and the coalition of all the men would benefit.

In a model with many to one matching, e.g. between firms and workers, or colleges and students, suppose that, in Example 1, m1 and m2 were in fact one agent, m12 with *two* positions. Then the example makes clear that if they pursue their true preferences, with a first choice of being matched to both w1 and w2, then they will end up matched to w2 and w3. But if they instead state their preferences so that they make their first proposals to w1 and w3, then the firm-proposing deferred acceptance algorithm will end after the first step, and agent m12 will be matched to w1 and w3, and will have profited from the misrepresentation, i.e. in this case even the conclusion of Theorem 3 will not carry over to a firm that employs many workers. But to make this precise, we have to consider a model in which firms or colleges such as m12, that wish to match to more than one worker or student, have preferences over groups and not just individuals.

2.2 More general models:

The two main directions in which related models and generalizations of the marriage model have developed into substantial literatures involve models in which price

setting is accomplished simultaneously with matching, and models of many-to-one matching. The developments I recount here build mostly on models of many-to-one matching in which prices are not explicit (but rather are reflected in the preferences that agents may have for jobs, in which e.g. wages are considered as part of the job description). However, I'll briefly mention some of the models intended to directly examine price setting, since they have become important in auction theory and design.

Models with money

One of the first questions that arose as economists began to think about matching models was how to incorporate money into the model, so that it could discuss not only matching, but also price determination. Shapley and Shubik (1972) took an important first step in this direction by considering an *assignment* model in which all preferences are monetary in nature. In their model, there are two sets of agents each of whom can be matched to an agent on the other side with the result that a certain monetary value will be produced, which may be freely reallocated.[6] They show that the core of the game is always nonempty, and that core outcomes do not involve any transfers between agents who are not matched to one another. The payoffs the players receive in the core (which are the dual solution to the linear program that finds the matching that maximizes the sum of the values) can be interpreted as the prices that one side receives from the other in the market, and are influenced by the general level of prices, so that, as in the marriage model, there are optimal points in the core for each side of the market, i.e. there is a point in the core at which all prices are simultaneously as low as possible, and another at which they are as high as possible.

A different approach was to look at the deferred acceptance algorithm itself as a form of auction mechanism, in which offers would include a price or wage (in this respect see particularly Demange, Gale, and Sotomayor 1986). Kelso and Crawford (1982) is the seminal paper in this regard (see also Crawford and Knoer, 1981). Kelso and Crawford explored a price adjustment process in which firms bid for workers' services by offering progressively higher wages in a firm-proposing deferred acceptance algorithm. In their model, firms could employ groups of workers, and so they also had to consider the form that firms' preferences over groups of workers could take that would allow the deferred acceptance algorithm to continue to produce a stable matching. They observed that this will be the case if workers occur in firms' preferences as substitutes rather than as complements, so that a firm that makes a set of offers to a group of workers and has some of those offers rejected does not wish to withdraw any offers to workers who have not rejected them. Formally, if a worker w is in a firm f's most favored choice set at a price p(f,w), then if some of the other offers in the choice set are rejected (and henceforth unavailable), worker w at price p(f,w) nonetheless remains in (one of) firm f's most preferred choice set among the remaining set of worker-price contracts available. (If not, then after firm f received a rejection from some worker w', it might also wish to withdraw its offer from worker w, who might have rejected another firm in order to hold

[6] Gary Becker (1981) famously explored related models to illuminate aspects of marriage and household production. See also Demange and Gale (1985) for a generalization of the assignment model in the spirit of GS.

firm f's offer, and so instabilities could result.) Hatfield and Milgrom (2005) present a masterful synthesis of much of auction and matching theory, extending this line of inquiry.[7]

Models of many-to-one matching

The college admissions model and its descendents without explicit price formation remain very useful for focusing on the matching aspects of models. The question that arose as economists began to think about using these models to study actual markets was how and to what extent the college admissions model outlined in GS could be extended to a well defined game involving many-to-one matching. A model of such a game would of course require the preferences of all the participants to be specified over possible outcomes, so colleges would need to have preferences defined not just over individual students, but over sets of students with whom they might be matched. If the model was to be a strategic game, colleges would also have strategies available that would not be available in a model of one-to-one matching.

Roth (1985) examined a 'minimal' extension of the GS college admissions model that would preserve the main properties of the marriage model to the extent possible, and allow strategic issues to be explored. To this end, consider a model consisting of colleges (that seek to match with multiple students) and students (who seek to match with a single college). In addition to the definitions given in GS, let a college C_i with quota q_i have preferences over sets of students with the property that, for any set S with $|S| < q_i$, and any students w and w', C_i prefers $S \cup w$ to $S \cup w'$ if and only if w is preferred to w' under college C_i's preferences over individual students. *Responsive* preferences defined in this way are a special case of preferences in which students are substitutes rather than complements in colleges' preferences (i.e. there is no student who is only desirable if some other student is also in the entering class).[8] This allows a marriage market to be constructed from any college admissions market by representing each college C_i by q_i "copies" each having C_i's preferences over individual students and seeking to be matched

[7] Indeed, for many years there were no good models that unified the parallel results not only for marriage and assignment models, but for more general auction and matching models. Principal among these were the dominant strategy results for the deferred acceptance algorithm and for second price auctions, which also give agents on one side of the market the core outcome most favorable to them. Aside from solidifying our understanding of the role of substitutes in preferences, Hatfield and Milgrom (2005) use Tarski's (1955) theorem (that an increasing function from a lattice to itself has a nonempty set of fixed points that form a sublattice) to unify these results, and connect them to the Lattice Theorem for the marriage model. (In this connection see also Adachi, 2003, Echenique and Oviedo 2004, 2006, Fleiner, 2003, and Ostrovsky, 2007, and, for different uses of Tarski's theorem in matching, Roth and Sotomayor 1988, and Sotomayor 2007.) Another kind of unification between the marriage and assignment models involves Abeledo and Rothblum's (1995) observation that the deferred acceptance algorithm is isomorphic to a dual simplex solution of a linear programming problem. There are still many open questions concerning when the core will be non-empty for preferences that aren't for substitutes, but see e.g. Echenique and Yenmez 2007, and Pycia (2007) for models in which workers care about not only which firm they are matched to but also who their colleagues are.

[8] Note that there may be many different preferences over groups of students that are responsive to the same preferences over individuals, since e.g. a college with responsive preferences over pairs of students could prefer to receive choices 1 and 4 to choices 2 and 3, or vice versa.

with one student (and such that all of C_i's copies appear consecutively in each student's preferences, wherever they originally ranked C_i). Because preferences are responsive, there is a one-to-one correspondence between the stable matchings of the original mant-to-one market and this related one-to-one marriage market. This allows Theorems 1 and 2 from GS to carry over immediately from the marriage model, and the extra structure of the derived marriage market even allows some results that concern only stable matchings to be strengthened.[9]

However, a correspondence between two sets of stable matchings does not help us draw conclusions that involve the comparison of a stable matching with an unstable one, and here we find results that do not generalize from the marriage model. In particular, we can formalize the re-analysis of Example 1 at the end of section 2.1 with the following formal results which show that the conclusions of Theorems 5, 8 and 9 above carry over from the marriage model only for the students, who continue to wish to match to a single college, and not to the colleges that wish to match to more than one student.

Theorem 10 (Roth, 1985) When the preferences over individuals are strict, the student-optimal stable matching is weakly Pareto optimal for the students, but the College-optimal stable matching need not be even weakly Pareto optimal for the colleges.

Theorem 11 (Roth, 1985) No stable matching mechanism exists which makes it a dominant strategy for all hospitals to state their true preferences, although the student-proposing deferred acceptance algorithm makes it a dominant strategy for all students to state their true preferences.

Of course, colleges with multiple seats to fill have more strategic options than just how they rank individual students, they also need to reveal how *many* students they wish to admit. Sönmez (1997) considered the strategic decisions faced by colleges constrained by some maximum capacity (and who in fact prefer to fill that capacity if they can get the students they prefer), but who were free to admit fewer students, and proved the following impossibility result.[10]

Theorem 12 (Sönmez, 1997) No stable matching mechanism makes it a dominant strategy for a college to always reveal its capacity.

3. Deferred acceptance algorithms in practice[11]

The first position taken by American doctors after receiving their M.D. degree is now called a residency, and it is a job that forms both the backbone of hospitals' labor force,

[9] We'll return to how some results for the marriage model can be strengthened when we discuss the Rural Hospital Theorem. See also Roth and Sotomayor (1989) in this connection.
[10] See Sonmez (1999) for some related results.
[11] The historical discussion of the American market for interns and residents is based on Roth (1984) and Roth (2003), in which more detailed accounts can be found, and on conversations in 2007 with Elliott Peranson, prompted by an initial draft of this section.

and a critical part of physicians' graduate education and career path. From 1900 to 1945, one avenue that hospitals' competition for new residents took was to try to hire residents earlier than other hospitals. This moved the date of appointment earlier, first slowly and then quickly, until by 1945 it was customary for residents to be hired almost two years before they would graduate from medical school and begin work. When I studied this in Roth (1984) it was the first market in which I had seen this kind of "unraveling" of appointment dates, but today we know that unraveling is a common and costly form of market failure. (Unraveling, in which offers become not only increasingly early, but also dispersed in time and of increasingly short duration, has occurred in other medical labor markets in the U.S., Canada, and Britain, and in numerous other kinds of markets; see Roth and Xing 1994 for an account of many unraveled markets).[12]

In 1945, the market for residents received an assist from a third party, the medical schools, who agreed not to release information about students before a specified date. This helped control the date of the market, but a new problem emerged: hospitals found that if some of the first offers they made were rejected, the candidates to whom they wished to make their next offers had often already accepted other positions. This led to exploding offers to which candidates had to reply immediately, before they could learn what other offers might be available, and to a chaotic market that shortened in duration from year to year, and resulted not only in missed agreements but also in broken ones. This kind of congestion also has since been seen in other markets, and in the extreme form it took in the American medical market by the late 1940's, it also constitutes a form of market failure (cf. Roth and Xing 1997, and Avery, Jolls, Roth, and Posner 2007 for detailed accounts of congestion in labor markets in psychology and law).

Faced with the costs of a market that was working very badly, the various American medical associations (of hospitals, students, and schools) agreed to employ a centralized clearinghouse to coordinate the market. After students had applied to residency programs and been interviewed, instead of having hospitals make individual offers to which students had to respond immediately, students and residency programs would instead be invited to submit rank order lists to indicate their preferences. That is, hospitals (residency programs) would rank the students they had interviewed, students would rank the hospitals (residency programs) they had interviewed, and a centralized clearinghouse—a matching mechanism—would be employed to produce a matching from the preference lists. Today this centralized clearinghouse is called the National Resident Matching Program (NRMP).

The original algorithm proposed for the medical clearinghouse (Mullen and Stalnaker, 1952) was an unstable mechanism that also made it risky, in an obvious way, for students to list their true preferences. It was replaced at the last minute with another algorithm, adapted from a prior regional clearinghouse called the Boston Pool Plan (see Roth, 2003),

[12] On the costs of such unraveling in some markets for which unusually good data have been available, see Niederle and Roth (2003) on the market for Gastroenterology Fellows, and Fréchette, Roth, and Unver (2007) on the market for post-season college football bowls. For some very recent unraveled markets, see Avery, Fairbanks, and Zeckhauser (2003) on college admissions; and Avery, Jolls, Posner, and Roth (2001) on appellate court clerks.

and this was the one actually employed starting in 1952, the first year that matches were decided through the clearinghouse. This change of algorithms was not well documented, however. Roth (1984) showed that the algorithm adopted in 1952 is equivalent to the hospital proposing deferred acceptance algorithm, and produces the hospital-optimal stable matching.[13]

Deferred acceptance algorithms seem to have been independently invented twice in connection with the NRMP medical clearinghouse. Elliott Peranson (personal communication) tells me that when he was first hired to propose a match like the NRMP match to be used for medical school admissions in 1972, he initially thought, based on the available literature, that the NRMP used the originally proposed algorithm as described by Mullen and Stalnaker (1952). He recognized its flaws, and independently developed a deferred acceptance algorithm.[14] After giving a presentation on this subject to the AAMC, he was invited to help organize the NRMP match, and implemented his version of a hospital proposing deferred acceptance algorithm.[15]

Indeed, deferred acceptance algorithms, with some variations, have been independently developed in quite a few markets over time, in part because, when a clearinghouse is called for to replace some market failure, deferred acceptance seems to correspond to a "folk model" of how markets proceed when they work in an orderly fashion.[16] See for example the rules for the telephone market by which clinical psychologists attempted to fill their positions (Roth and Xing 1997), prior to adopting a centralized clearinghouse in 1999. That market was codified by rules that looked very much like the deferred acceptance algorithm: employers made offers starting at 9:00am on "selection day", and candidates were obliged to hold no more than one offer at a time; they were to immediately call and reject any offer they were holding if they decided to hold a subsequent offer. Aside from the fact that this was a decentralized market (in which

[13] Although the internal operation of the algorithm was organized quite differently, Roth (1984) showed that it is equivalent, in the sense that it always produced the hospital optimal stable matching in terms of the stated preference lists. (The GS deferred acceptance algorithm produces a stable match from any stated preferences, the Boston Pool algorithm as adopted in 1952 only does so if preference lists are such that any student who is listed by some hospital also lists that hospital; consequently it required an initial step in which submitted preferences were edited if necessary to have this property.)

[14] "The night before we were leaving for Washington to give the seminar I was working through an example I was going to present, and realized the algorithm wasn't working very well. I called my colleague, and together over a couple of beers we came up with what turned out to be the deferred acceptance algorithm. I revised the seminar late that night to say we didn't think NRMP worked properly." (personal communication, 3/30/07)

[15] In the process of doing so, he also discovered his mistake about the algorithm that was in fact being used. Peranson writes: "...we went to the NRMP office, I gave them some examples to work through to see if I was right about how their algorithm didn't work... The next day, we went back to NRMP and, based on their solutions to the examples, I discovered that they were in fact using a deferred acceptance algorithm. (At least they claimed that's how it worked - it had been computerized just recently, the computer system didn't work very well, and I'm not sure anyone knew what the computer programs really did to produce a matching)."

[16] I include here clearinghouses that are equivalent to deferred acceptance algorithms even if this equivalence is not always apparent. See e.g. Guillen and Kesten (2007) who discuss study a procedure used to allocate rooms in an MIT dormitory, which they show is equivalent to a deferred acceptance algorithm with certain conventions applied to the "preferences" that rooms have for students.

employers and candidates could make decisions at each part of the process, rather than submitting a rank order list of preferences to be followed from the outset), it resembled the deferred acceptance algorithm except for one crucial difference: the process terminated not when no further offers remained to be made, but rather at 4:00pm, at which point offers that were not yet accepted would no longer be available. Because this didn't allow all offers to be made, a great deal of strategic behavior had developed that complicated the operation of the market, and eventually led to the adoption of a centralized clearinghouse.[17]

Over the years, changes in the medical marketplace made a simple deferred acceptance algorithm less able to accommodate the requirements of the market. These included some which preserved preferences as substitutes, and some that did not. An example of the first kind of change is that hospitals sometimes wish to manage the number of residents in multiple programs who might do similar kinds of work, so that they might want to have 10 residents in program A and 5 in program B, but be sure to have 15 in total, so that if program B filled only 4 of its 5 positions they would like to fill 11 positions in program A. Note that the hospital's preferences to hire an 11^{th} person in program A only if it cannot hire a 5^{th} person in program B treat those people as substitutes, rather than complements, e.g. those preferences would not cause a hospital to withdraw an offer it had made when it received a rejection of another offer.

An example of a change in the market that involved preferences that are not substitutes had to do with the growing number of married couples graduating from American medical schools and wishing to be matched to jobs in the same vicinity. This hadn't been a problem when the match was created in the 1950's, when virtually all medical students were men. But by the 1960's it was enough of a problem that the match had tried to accommodate couples. In the early 1970's the first attempt to fully automate the matching of couples permitted two students to identify themselves as a couple, and define "communities" of hospitals in which they would like to both be matched. However, both members of the couple continued to submit rank order lists as individuals, and had to declare one of their members the "leading member," with the preference list of the other member then being processed in a manner to facilitate matching to the same community. While this mechanism often matched the members of the couple to positions in the same city, couples often declined to take these positions, and instead managed to arrange positions they preferred. In Roth (1984) I observed that if couples were modeled as having preferences over *pairs* of positions, this could be understood as responding in a

[17] In the 20 years before our study of that market it had regularly undergone rule changes designed to address some of the failings of the market. Most notable of these was that the duration of the market had *decreased* from a week to a day. This allowed everyone to stay by their phones from beginning to end, and, by reducing the time it took to reject an offer after receiving a preferable one, it increased the effective length of the market, basically the number of offers that could be made. But, as we argued in Roth and Xing (1997), there wasn't enough time for the market to fully clear in a straightforward way, because after a short time only a very few people were holding two offers at a time, and even very speedy phone calls didn't allow the deferred acceptance process to reach its natural conclusion; the deadline was always potentially binding. (For this reason, this was a market in which employers elicited from applicants the promise that they would accept a job immediately if called on the morning of selection day, and often made offers to applicants from whom they had received such promises.)

natural way to the instability of the resulting matching.[18] I further observed that when couples had preferences over pairs of positions, the set of stable matchings could be empty. And, secure in my observer status as an economic theorist (and not yet a market designer), I concluded by noting, therefore, "that even when couples are allowed to state their preferences more accurately, the problem of finding a stable outcome may still be intractable..."

I therefore felt at least a little dismay when, in 1995, I was invited to direct the redesign of the medical match, in response to a crisis in confidence that had developed regarding its ability to continue to serve the medical market, and whether it appropriately served student interests.[19] As a designer, I would no longer have the luxury of simply observing that some problems are hard: e.g. more than a thousand people go through the match each year as members of couples, and have to be matched.[20] And the empirical evidence was clear that stability was important to the success of matches. Roth (1990, 1991) had studied the clearinghouses that had been tried in the various regions of the British National Health Service, after those markets unraveled in the 1960's. A Royal Commission had recommended that clearinghouses be established on the American model, but since the American medical literature didn't describe in detail how the clearinghouse worked, each region of the NHS adopted a different algorithm for turning rank order lists into matches, and the unstable mechanisms had largely failed and been abandoned, while the stable mechanisms succeeded and survived.[21]

Roth and Peranson (1999) reports on the new clearinghouse algorithm that we designed, a generalized applicant-proposing deferred acceptance algorithm that allows couples to state preferences over pairs of positions, allows hospitals to specify reversions of positions, and deals with some other features of the market as well. Unlike the deferred acceptance algorithm for simpler matching markets, this does not produce a stable matching on the first pass through agents' preferences. Instead, a list of potential blocking pairs is accumulated (e.g. every time one member of a couple is displaced from a position and causes the other member of the couple to be withdrawn from a position), and the algorithm attempts to satisfy these pairs one at a time, using a class of algorithms explored in Roth and Vande Vate (1990). But, to be clear about the relationship to the

[18] Couples are of course subject to the iron law of marriage, that you can't be happier than your spouse (at least not in equilibrium).

[19] In this respect, the fact that the algorithm was a hospital proposing deferred acceptance algorithm became a point of debate, as did the question of whether students were or should be submitting preferences strategically. By this time many adaptations to the initial algorithm had already been made, including allowing couples to state preferences over pairs of positions, so it was no longer a simple hospital-proposing deferred acceptance algorithm.

[20] I was acutely aware that, of all the results on matching presented in Roth and Sotomayor (1990), the only ones that directly applied to the medical match were the counterexamples such as those concerning couples and the possible absence of stable matchings.

[21] Of course, there are other differences between regions of the British health service than how they organized their medical clearinghouses, so there was also room for controlled experiments in the laboratory on the effects of stable and unstable clearinghouses; see Kagel and Roth (2002). And some of the differences between British and U.S. markets raised interesting theoretical questions as well, see e.g. Roth (1991) and Irving (1998) on various consequences of the fact that graduating British doctors need two positions.

deferred acceptance algorithm in GS, this is a generalization in the sense that if all hospitals and applicants submitted simple preferences (without couples, reversions of positions, etc.) the algorithm would produce the applicant optimal stable matching. The new algorithm has been used by the NRMP since 1998, and has subsequently been adopted by over three dozen labor market clearinghouses (see Table 1).[22] The empirical evidence that has developed in use is that the set of stable matchings is very seldom empty but is often very small, and in the next section we'll discuss some of the new theoretical questions these results raise.

An interesting historical note is that the use of deferred acceptance algorithms has been explicitly recognized as part of a pro-competitive market mechanism in American law. This came about because in 2002, over a dozen law firms representing three former medical residents brought an antitrust suit challenging the use of the matching system for medical residents. The theory of the suit was that the matching system was a conspiracy to hold down wages for residents and fellows, in violation of the Sherman Antitrust Act.[23] Niederle and Roth (2003) observed that, empirically, the wages of medical specialties with and without centralized matching in fact do not differ. An interesting, game theoretic aspect of the suit itself was that it was a double class action suit, that sought not only to represent the class of all former medical residents, but sought to sue a class of defendants that included all hospitals that employ residents, including several dozen named defendants. This had the effect of making the legal defense costs very high, as lawyers for the many defendants had to coordinate with one another in each of the many preliminary motions and hearings expected to stretch over years. Thus the strategy of the plaintiffs seemed aimed at trying to force a financial settlement before the otherwise lengthy process reached trial. However the same strategy that made legal defense costly made it possible for the defendants to seek legislative relief from the suit, since virtually

[22] The establishment of a clearinghouse interacts with the extent of a market. Thus in Table 1, almost all entry level medical residencies, in almost all specialties, are handled simultaneously by the NRMP clearinghouse, and we can therefore consider that to be a single market, even though very few if any applicants will be in the market for, say, both psychiatry and surgery positions. On the other hand, the various subspecialties of Internal Medicine are separate markets, partly due to the different elective training that residents must take to prepare for them, but partly now because they have separate matches. Following our work in helping to re-establish a Gastroenterology match (Niederle and Roth, 2004, 2005, Niederle, Proctor and Roth, 2006), Muriel Niederle and I are now discussing with Orthopedic Surgery fellowship directors the possibility of establishing a comprehensive match, and this involves assessing the extent to which e.g. Sports Medicine and Spine Surgery do or could constitute separate markets, or parts of a single market for orthopedic surgery fellowships.

[23] Bulow and Levin (2006) sketch a simple model of one-to-one matching in which a centralized clearinghouse, by enforcing impersonal wages (i.e. the same wage for any successful applicant) could cause downward pressure on wages (see also Kamecke 1998). Subsequent analysis suggests more skepticism about any downward wage effects in actual markets. See, for example, Kojima (forthcoming) which shows that the Bulow-Levin results don't follow in a model in which hospitals can employ more than one worker, and Niederle (forthcoming) who shows that the results don't follow in a model that includes reversions. (Niederle shows how, in a model with reversions, like those in the NRMP, the difference between hospital and applicant proposing deferred acceptance algorithms may be different than in a simple matching market, e.g. in some circumstances both the hospital and student-proposing versions of the algorithm could produce the hospital optimal stable matching even when it is different from the student optimal stable matching.) Crawford (forthcoming) considers how the deferred acceptance algorithm of Kelso and Crawford (1982) could be adapted to adjust personal wages in a centralized clearinghouse.

every senator and congressman has a hospital in his district. The struggle thus shifted to the Congress, and Public Law 108-218 (2004) notes that the medical match is a pro-competitive market mechanism, not a conspiracy in restraint of trade.[24]

A different area of application developed during this period, involving the use of clearinghouses based on deferred acceptance algorithms for assigning children to schools. The clearinghouses designed for New York City high schools and for Boston Public Schools at all levels are described in Abdulkadiroğlu, Pathak, and Roth (2005, 2007) and Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005, 2007). The New York City clearinghouse has been in operation since 2003 (for students entering high school in September 2004), and the Boston clearinghouse went into operation in 2005-06 for students entering school in September 2006. The problems that these designs are intended

---

[24] The law states in part: "Congress makes the following findings: For over 50 years, most United States medical school seniors and the large majority of graduate medical education programs (popularly known as `residency programs') have chosen to use a matching program to match medical students with residency programs to which they have applied. These matching programs have been an integral part of an educational system that has produced the finest physicians and medical researchers in the world.

"Before such matching programs were instituted, medical students often felt pressure, at an unreasonably early stage of their medical education, to seek admission to, and accept offers from, residency programs. As a result, medical students often made binding commitments before they were in a position to make an informed decision about a medical specialty or a residency program and before residency programs could make an informed assessment of students' qualifications. This situation was inefficient, chaotic, and unfair and it often led to placements that did not serve the interests of either medical students or residency programs.

"The original matching program, now operated by the independent non-profit National Resident Matching Program and popularly known as `the Match', was developed and implemented more than 50 years ago in response to widespread student complaints about the prior process. This Program includes on its board of directors individuals nominated by medical student organizations as well as by major medical education and hospital associations.

"The Match uses a computerized mathematical algorithm, as students had recommended, to analyze the preferences of students and residency programs and match students with their highest preferences from among the available positions in residency programs that listed them. Students thus obtain a residency position in the most highly ranked program on their list that has ranked them sufficiently high among its preferences. Each year, about 85 percent of participating United States medical students secure a place in one of their top 3 residency program choices.

"Antitrust lawsuits challenging the matching process, regardless of their merit or lack thereof, have the potential to undermine this highly efficient, pro-competitive, and long-standing process. The costs of defending such litigation would divert the scarce resources of our country's teaching hospitals and medical schools from their crucial missions of patient care, physician training, and medical research. In addition, such costs may lead to abandonment of the matching process, which has effectively served the interests of medical students, teaching hospitals, and patients for over half a century.

"… It is the purpose of this section to—confirm that the antitrust laws do not prohibit sponsoring, conducting, or participating in a graduate medical education residency matching program, or agreeing to do so; and ensure that those who sponsor, conduct or participate in such matching programs are not subjected to the burden and expense of defending against litigation that challenges such matching programs under the antitrust laws."

to solve are quite different: in New York, the school choice system is a two-sided market in which the schools are active players, but in Boston the schools are passive, and the problem is a one-sided market as studied by Abdulkadiroğlu and Sönmez (2002), who observed that deferred acceptance algorithms might often be sensible even when stability *per se* was not required to stop blocking pairs from forming. In Boston in particular, where the new clearinghouse is a straightforward student-proposing deferred acceptance algorithm, the central issue in its adoption was that it would give students and their families a dominant strategy to state their true preferences. School choice in Boston had previously been organized through a centralized clearinghouse that was not at all strategy proof[25], and one of the factors in the adoption of the new system was that it would eliminate the need for students and families to "game the system."

These school choice clearinghouses are the applications that are most in the spirit of the "college admissions" discussed by GS, but they also raise new theoretical questions. The one I'll address in the next section has to do with how indifferences are treated, which is an important question in a system in which many children may be indistinguishable from the point of view of a given school, but in which ties must somehow be broken because there are insufficient places at particular schools to accommodate all children in the same indifference class.

4. New theory, open questions and research directions

In this section, I'll attempt to briefly and fairly informally describe some of the new theoretical questions that arose in the course of analyzing and designing these various clearinghouses.

One question that arose in connection with the resident match concerns rural hospitals. These often cannot fill all their residency positions, and a disproportionate number of those they do hire are graduates of foreign medical schools. Since the vast majority of residency positions are filled through the match, the question was, when a new match algorithm is being written, can it relieve the plight of these hospitals? Given that the empirical evidence supported the view that stability is an important component of match success, the following theorem answers that question in the negative. (Note that the first part of the theorem is a straightforward generalization of the result of McVitie and Wilson for the marriage model, but the second part has no parallel when matching is one-to-one.)

Theorem 13: (Rural Hospital Theorem, Roth 1986): When all preferences over individuals are strict, and hospitals have responsive preferences, the set of students employed and positions filled is the same at every stable matching. Furthermore, any hospital that has some empty positions at some stable matching is assigned precisely the same set of students at every stable matching.

[25] A great deal can be learned by studying flawed existing allocation systems. In this connection see e.g. Balinski and Sonmez (1999), Ergin and Sonmez (2006), and Pathak and Sonmez (2007).

As already mentioned, a different kind of problem was caused by the presence of couples, which might make the set of stable matchings empty. The counterexamples in Roth (1984) and also in unpublished work by Sotomayor (see the Appendix) each involved two couples and four hospitals, in which blocking pairs had to be identified for 24 matchings to show that none were stable. Today we can show this much more simply. Klaus and Klijn (2005) show that, except when all couples have preferences that are ("weakly") responsive to their individual preferences (i.e. except in the case of "couples" who have preferences as if they were single, and don't alter their preferences for their own jobs depending on where the other member of the couple is employed), examples that include a given couple can always be constructed with no stable matchings.[26]

Example 2--market with one couple and no stable matchings (Klaus and Klijn): Let c=(s1,s2) be a couple, and suppose there is another single student s3, and two hospitals h1 and h2. Suppose that the acceptable matches for each agent, in order of preference, are given by

c: (h1,h2)[27]; s3: h1, h2,
h1: s1, s3; h2: s3, s2

Then no individually rational matching μ (i.e. no μ that matches agents only to acceptable mates) is stable. We consider two cases, depending on whether the couple is matched or unmatched.

Case 1: $\mu(c)$=(h1,h2). Then s3 is unmatched, and s/he and h2 can block μ, because h2 prefers s3 to μ (h2)=s2.

Case 2: μ (c)=c (unmatched). If μ (s3)=h1, then (c, h1,h2) blocks μ. If μ (s3)=h2 or μ (s3)=s3 (unmatched), then (s3,h1) blocks μ.

Aldershof and Carducci (1996) further showed by example that, even when the set of stable matchings is nonempty, when couples are present there need not be a lattice structure, or optimal stable matchings, nor is the set of unmatched agents the same at every stable matching. Indeed, there aren't any ready parallels to any of the standard results for models in which couples are present (with the exception of the various impossibility results; the presence of couples certainly doesn't make it easier to construct strategy-proof stable mechanisms).

---

[26] In keeping with the general informality of this review, I don't include a formal definition of instability involving couples, who e.g. could participate in blocking pairs together with one or two employers. The example should nevertheless be clear.

[27] Couple c submits a preference list over pairs of positions, and specifies that only a single pair, h1 for student s1 and h2 for student s2 is acceptable. Otherwise couple c prefers to remain unmatched. Note that this preference is not responsive to any individual preferences that has, say, s1 preferring h1 to being unmatched, since the couple does not prefer the option at which s1 is matched to h1 and s2 is unmatched to the option of both being unmatched. For a couple, this could make perfect sense, if e.g. h1 and h2 are in a different city than the couple now resides, and they will move only if they find two good jobs.

However, one of the empirical observations of Roth and Peranson (1998) was that, in the large medical market studied there, with over 20,000 positions and almost that many applicants, the standard results seemed to provide a very good approximation. For example, only about one in 100,000 positions switched from being matched to unmatched when applicant and employer-proposing variants of the algorithm were compared. In general, the set of stable matchings seemed to be quite small, and there appeared to be very few opportunities for profitable manipulation of preferences or capacities, even for agents who were not on the proposing side. This latter result is surprising, because in a simple marriage model in which agents on each side have preferences over all the agents on the other side, the set of stable matchings becomes large as the number of agents grows (e.g. when preferences are uncorrelated), so that at any stable matching, many agents are not receiving their optimal stable match, and so could profitably manipulate their preferences (recall Theorem 7). But in the medical market, most rank order lists submitted by applicants had no more than fifteen residency programs listed. Roth and Peranson showed computationally that if, as such a market gets large, the number of places that a given applicant interviews (and hence the size of his rank order list) does not grow, then the set of stable matchings becomes small.

This observation has sparked some new theory, and raised some conjectures. Immorlica and Mahdian (2005) showed analytically that in the marriage model with uncorrelated preferences, the set of people who are matched to different mates at different stable matchings grows small as the market grows large in this way, and that therefore the opportunities for profitable manipulation grow small.

Kojima and Pathak (2007) substantially extend this result to the case of many to one matching, in which opportunities for employers to profitably manipulate can occur even when there is a unique stable matching, and in which employers can manipulate capacities as well as preferences. They look at a sequence of markets with randomly generated preferences (in which employers may have arbitrary responsive preferences), indexed by the number of employers. They define an infinite sequence of markets of increasing size n to be *regular* if the length k of applicants' random preference lists remains constant[28], the capacity of each employer is bounded, the number of applicants doesn't grow faster than the total capacity of all employers, and each applicant is acceptable to every employer. They prove the following.

Theorem 14 (Kojima and Pathak, 2007): In the limit, as n goes to infinity in a regular sequence of random markets, the proportion of employers who might profit from (any combination of) preference or capacity manipulation goes to zero in the worker proposing deferred acceptance algorithm.

A second empirical observation made in the resident match data, and in the other matches in Table 1, is that, even when couples are present, it is a very rare occurrence for the set of stable matchings to be empty. In the vast majority of cases a stable matching is

[28] Students' preferences are drawn independently from the same distribution. For each student, colleges are selected from the same probability distribution, one at a time, until k colleges have been picked (for a fixed k), these are the student's k top choices in the order selected.

produced.[29] This is true even though we have seen that it is easy to produce small examples in which no stable matching exists. An open question is why this is so. I offer the following loose conjecture:

Conjecture: In the limit, as n goes to infinity in a regular sequence of random markets in which the proportion of couples is bounded [or, goes to zero], the probability that the set of stable matchings is empty goes to zero.

This conjecture is loosely specified in a number of ways, one of which is that I haven't specified from what kind of distribution couples' preferences should be drawn. We already know from Klaus and Klijn (2005) that if the couples all look too much like individuals, the set of stable matchings will always be nonempty. I would imagine that a simple way to start would be to consider a world in which there are multiple locations, with couples having some nontrivial tradeoff between personal preferences for positions, and the desire to be in the same location as their spouse.

Note that this conjecture might seem counterintuitive, given that the set of stable matchings becomes small as the market becomes large in this way. That is, it might be thought that in larger markets, with smaller sets of stable matchings, the likelihood that the set of stable matchings would be empty would be larger. My conjecture, based on field observations, is the opposite.

As noted earlier, the question of indifferences arises when considering school choice. When schools, say, have responsive preferences arising from non-strict preferences over individuals, stable matchings will still exist, but optimal stable matches in the strong sense of Theorem 2 will no longer exist. Instead, we'll have to consider the set of stable matchings that are Pareto optimal for students, i.e. each with the property that there is no stable matching that gives some students more preferred matches without giving other students less preferred matchings. The problem is how to break ties to achieve one of these Pareto optimal matchings, since random or arbitrary tie breaking may fail to do so (even in the simple case of one-to-one matching). We can see this by looking back at Example 1, and modifying it so that not all preferences are strict.

Example 1' (Tie breaking can be inefficient): All preferences are as in Example 1, except for those of w1, and let w1 be indifferent between any of m1, m2, m3. Suppose, at step 0 of the deferred acceptance algorithm, the ties in w1's preferences are broken so as to produce the (artificial) strict preference P(w1) = m1, m2, m3. Now the analysis is the same as in Example 1, and the deferred acceptance algorithm operating on the artificial strict preferences produces μ_M **= [(m1,w1); (m2,w3); (m3,w2)]**, at which m1 and m3 each receive their second choice mate (while m2 receives his last choice). As in our original analysis of Example 1, matching μ **= [(m1,w2); (m2,w3); (m3,w1)]**, is Pareto superior for the men, as m1 and m3 each receive their first choice, so they are both strictly better off than at μ_M, and m2 is not worse off. The difference from Example 1 is

[29] The Roth-Peranson algorithm could fail to find a stable matching even when one exists. In the matches in Table 1, my understanding is that in the last ten years there have been fewer than half a dozen occasions on which no stable matching was found.

that in the present example, since w1 is in fact indifferent between her possible mates, the matching μ is also stable. The pair (w1, m2) is not a blocking pair for μ, and only appeared to be in the deferred acceptance algorithm because of the arbitrary ways in which ties were broken to make w1's preferences look strict.

So, there are costs to arbitrary or random tie breaking. Kestin (2004), Erdil and Ergin (2007, 2006) and Abdulkadiroğlu, Pathak, and Roth (2007) each explore this from different angles.[30]

Kestin observes that the reason that students are collectively better off at μ than at μ_M (in Example 1 as well as 1')[31] can be traced to the operation of the deferred acceptance algorithm, in which m2's fruitless attempt to match with w1 harms m1 and m3 without yielding m2 any benefit. Kestin defines an *efficiency adjusted deferred acceptance mechanism* that produces μ in Examples 1 or 1' by disallowing the blocking pair (w1, m2) even in Example 1 with strict preferences, under an extended notion of "reasonable fairness" that generalizes stable matchings. He further observes that there is no mechanism that is Pareto efficient, reasonably fair, and strategy proof.

To understand Erdil and Ergin's approach, note that in Example 1', the Pareto improvement from μ_M to μ arises from an exchange of positions between m1 and m3. The key observation is that this exchange doesn't introduce any new blocking pairs, since, among those who would like to change their positions, m1 and m3 are among the most preferred candidates of w1 and w2. Since there weren't any blocking pairs to the initial matching, this kind of exchange can occur without any new blocking pairs being created by the exchange of positions.

Formally, Erdil and Ergin define a *stable improvement cycle* starting from some stable matching to be a cycle of students who each prefer the school that the next student in the cycle is matched to, and each of whom is one of the school's most preferred candidates among the students who prefer that school to their current match. They prove the following theorem.

Theorem 15 (Erdil and Ergin, 2007): If μ is a stable matching that is Pareto dominated (from the point of view of students) by another stable matching, then there is a stable improvement cycle starting from μ.

Erdil and Ergin show that, as a result, there is a computationally efficient algorithm that produces stable matchings that are Pareto optimal with respect to students. The initial step of the algorithm is a student-proposing deferred acceptance algorithm with arbitrary tie-breaking of non-strict preferences by schools. The output of this process (i.e. the

[30] In the computer science literature there has been a focus on the *computational* costs of non-strict preferences, which adds to the computational complexity of various calculations, see e.g. Irving, Manlove, and Scott 2000.)

[31] Kestin's concern is not with tie-breaking *per se* but with the general question of Pareto efficiency for students in school choice problems in which schools don't have preferences that enter into welfare or strategy considerations.

student optimal stable matching of the market with artificially strict preferences) is then improved by finding and satisfying stable improvement cycles, until no more remain. They observe, however, that this algorithm is not strategy-proof, i.e. unlike the student-proposing deferred acceptance algorithm, this deferred acceptance plus stable improvement cycle algorithm doesn't make it a dominant strategy for students to state their true preferences. They show in fact that no mechanism that always produces a stable matching that is Pareto optimal for the students can be strategy proof.

Abdulkadiroğlu, Pathak and Roth (2007) show that no mechanism (stable or not, and Pareto optimal or not) that is better for students than the student proposing deferred acceptance algorithm with tie breaking can be strategy proof. That is, we'll say that one mechanism *dominates* another if for every profile of preferences the first mechanism produces a matching that is at least as good for every student as the matching produced by the second mechanism, and for some preference profiles the first mechanism produces a matching that is preferred by some students. Following the design of the New York and Boston school choice mechanisms, define a *tie-breaking rule* T to be an ordering of students that is applied to any school's preferences to produce a strict order of students within each of the school's indifference classes (that is, when a school is indifferent between two students, the tie breaking rule determines which is preferred in the school's artificial strict preferences). *Deferred acceptance with tie breaking rule T* is then simply the deferred acceptance algorithm operating on the strict preferences that result when T is applied to schools' preferences.

Theorem 16 (Abdulkadiroğlu, Pathak and Roth, 2007): For any tie-breaking rule T, there is no mechanism that is strategy proof for every student and that dominates student proposing deferred acceptance with tie-breaking rule T.

An interesting open problem has to do with the extent to which actual welfare gains could be achieved using a non-strategy-proof mechanism. Using the preference data submitted in a recent year by students in the New York City high school choice process, Abdulkadiroğlu, Pathak and Roth (2007) show that as many as 6,000 students (out of around (90,000) could be made better off through stable improvement cycles. Of course, this is an estimate made based on the preferences submitted in the current school choice mechanism, based on a student proposing deferred acceptance mechanism with tie breaking. The efficiency losses are only identifiable because the submitted preferences can be taken as a reasonable approximation of true preferences, since they were elicited from an algorithm that is strategy proof for the large majority of students (recall that the NYC implementation is only approximately a deferred acceptance algorithm). Nothing is now known about what kinds of preferences might be submitted (e.g. at equilibrium) of an algorithm that attempted to achieve the potential efficiency gains.

In turn, there would be costs to losing strategy-proofness. A strategy proof mechanism allows school officials to give simple advice to families about how to participate in the matching system. And it yields valuable preference data. (In New York City, schools revealed to be unpopular through the preference data have been closed.)

So an interesting theoretical area with important practical implications will be to explore the equilibrium, incentive, and efficiency properties of mechanisms that seek to achieve efficiency gains over deferred acceptance algorithms, in matching markets in which indifferences are important. The properties of such mechanisms will be interesting both in general, and over limited domains of preferences on which results might differ. (For example, in the Boston schools preference data, there seems to be virtually no scope for student welfare gains in comparison to the student optimal stable matching.)

Note that Theorem 16 doesn't say that there aren't any strategy proof mechanisms that are Pareto optimal for students (it only says that any such mechanism cannot always be better for students than deferred acceptance with tie breaking). Indeed, in the design of the Boston school choice system, in which stability of the final matching was not of primary importance (as it was in NYC), we considered another mechanism, based on the *top trading cycles* mechanism, attributed to David Gale in Shapley and Scarf (1974). That mechanism is strategy proof for students (Roth, 1982), and can be generalized in a way that potentially also would have made sense for Boston schools (Abdulkadiroğlu and Sönmez, 1999, 2003). But that is a story for another occasion.[32]

5. Matching and Market Design:

Although I have focused in this review on the deferred acceptance algorithm, I will end by briefly indicating some of the larger questions in economics that are illuminated by studying the kinds of markets in which deferred acceptance clearinghouses are used. A lot has been learned (and much more remains to be learned) about the detailed game theoretic microstructure of markets by trying to answer these questions. Why have clearinghouses based on the deferred acceptance algorithm (as well as other kinds of clearinghouses) been of use in markets? What are the kinds of market failure for which such a clearinghouse is a solution? To put it another way, what do we learn about markets, and about clearinghouses organized around deferred acceptance algorithms, from the markets that have employed such clearinghouses?

Muriel Niederle and I have gotten an unusually clear view of this by studying the market for Gastroenterology fellows, and helping gastroenterologists reorganize it.[33] Like the

[32] The top trading cycles (TTC) algorithm forms the basis of a new appeals process in the NYC high school choice system. It also served as an initial conceptual basis for the algorithm organizing the New England Program for Kidney Exchange (Roth, Sonmez, Unver 2004, 2005), although various considerations have since led in other directions (see e.g. Roth, Sönmez, Unver 2007, and Abraham, Blum, and Sandholm 2007). Like the deferred acceptance algorithm, TTC has also been the source of elegant theory (cf. Abdulkadiroglu and Sönmez 1998 and Sönmez and Unver, 2005), which connects it to the broader subject of general allocation mechanisms for indivisible goods (see e.g. Ergin 2002, Ehlers and Klaus 2006, Kesten 2006, Papai 2000). The behavior of TTC with respect to tie-breaking when preferences are not strict is surprisingly different from that of deferred acceptance: Pathak (2007) shows that it is insensitive to how ties are broken.

[33] The market for gastroenterology fellows was organized through a centralized fellowship match, the Medical Specialties Matching Program (MSMP organized by the NRMP) from 1986 to the mid nineties. The arrangement fell apart (see McKinney, Niederle and Roth, 2005 for an investigation into the causes of this collapse), and for the next decade the market operated in a decentralized way. It re-established a match in 2006, see Niederle and Roth (2007).

market for medical residents before it, this is a market with a complex history, which in recent years experienced periods of

- *unraveling* of appointment dates so that offers were being made at early and dispersed time;
- *congestion* in the making of offers, so that employers found that there was insufficient time to make all offers they would like in a timely way. (By the time a candidate had rejected an offer, the next choice candidate may already have accepted an offer elsewhere.)
- *lots of strategic play* in terms of the scheduling of interviews and offers, as it became unsafe for participants to act straightforwardly on their preferences.

To summarize, like a number of other markets, the gastroenterology fellows market experienced problems in maintaining a thick market, dealing with congestion when the market is thick, and in general making it strategically simple to participate in the market.

These problems, together or separately, are often the reasons that markets adopt a clearinghouse. By making offers through an algorithm, congestion problems can be solved, as algorithms operate very fast. Furthermore, using a deferred acceptance algorithm makes it safe for the proposing side of the market to make offers according to their preferences only, and, if the market is large, this is true for both sides of the market. Therefore if there is sufficient participation in the centralized clearinghouse, the market is thick, uncongested, and safe even for nonstrategic participants.

Of course, even markets that are working inefficiently may be at an equilibrium that makes it difficult to move them to another, better equilibrium. In the case of the gastroenterology market, it had become customary for fellowship programs to hire internal medicine residents up to two years before they would begin employment, and there was interest in establishing a clearinghouse that would operate only one year before employment would begin. But even though most fellowship programs were eager to make this transition, there was widespread concern that it would be disadvantageous to do so if other, competing fellowship programs continued to hire earlier than the clearinghouse. To make it safe for programs to wait a year and participate in the match, the four gastroenterology professional organizations eventually adopted a resolution, proposed in Niederle, Proctor and Roth (2006), that allowed applicants who had accepted offers before the match to later change their mind and decline them if they wanted to participate in the match. This changed the equilibrium in the decentralized market, and allowed participants to confidently plan to participate in the first run of the clearinghouse in June of 2006, for fellows who would begin one year later.

As this illustrates, while much of market design has so far concentrated on market failures that can be fixed with centralized clearinghouses of one sort or another (see e.g. Milgrom 2004 in this connection, and Roth 2002), there is often a decentralized market or potential market in the background that must also be understood. One promising area of

research will be to better understand how the problems of providing thickness, dealing with congestion, and making markets safe are dealt with in decentralized markets.[34]

6. Conclusion:

Gale and Shapley (1962) initiated and advanced what has grown into not only a substantial academic literature on matching models and mechanisms, but also an emerging empirical literature on market failure and market organization, and a growing "economic engineering" practice of market design.[35]

[34] For some studies that focus on aspects of decentralized markets such as unraveling, see Damiano, Li and Suen (2005), Li and Rosen (1998), Li and Suen (2000, 2004), Niederle and Roth (2007), Niederle and Yaariv (2007), and Unver (2001).

[35] Nobel committee, take note.

Table 1: Labor markets that adopted the Roth-Peranson clearinghouse design after 1998 (and date of first use of a centralized clearinghouse of some sort):

- Postdoctoral Dental Residencies in the United States
  - Oral and Maxillofacial Surgery (1985)
  - General Practice Residency (1986)
  - Advanced Education in General Dentistry (1986)
  - Pediatric Dentistry (1989)
  - Orthodontics (1996)
- Psychology Internships in the United States and Canada (1999)
- Neuropsychology Residencies in the U.S. and Canada (2001)
- Osteopathic Internships in the United States (before 1995)
- Pharmacy Practice Residencies in the United States (before 1994)
- Articling Positions with Law Firms in Alberta, Canada (1993)
- Medical Residencies in the United States (NRMP) (1952)
- Medical Residencies in Canada (CaRMS) (before 1970)
- Specialty Matching Services (SMS/NRMP):
  - Abdominal Transplant Surgery (2005)
  - Child & Adolescent Psychiatry (1995)
  - Colon & Rectal Surgery (1984)
  - Combined Musculoskeletal Matching Program (CMMP)
    - Hand Surgery (1990)
  - Medical Specialties Matching Program (MSMP)
    - Cardiovascular Disease (1986)
    - Gastroenterology (1986-1999; rejoined in 2006)
    - Hematology (2006)
    - Hematology/Oncology (2006)
    - Infectious Disease (1986-1990; rejoined in 1994)
    - Oncology (2006)
    - Pulmonary and Critical Medicine (1986)
    - Rheumatology (2005)
  - Minimally Invasive and Gastrointestinal Surgery (2003)
  - Obstetrics/Gynecology
    - Reproductive Endocrinology (1991)
    - Gynecologic Oncology (1993)
    - Maternal-Fetal Medicine (1994)
    - Female Pelvic Medicine & Reconstructive Surgery (2001)
  - Ophthalmic Plastic & Reconstructive Surgery (1991)
  - Pediatric Cardiology (1999)
  - Pediatric Critical Care Medicine (2000)
  - Pediatric Emergency Medicine (1994)
  - Pediatric Hematology/Oncology (2001)
  - Pediatric Rheumatology (2004)
  - Pediatric Surgery (1992)
  - Primary Care Sports Medicine (1994)
  - Radiology
    - Interventional Radiology (2002)
    - Neuroradiology (2001)
    - Pediatric Radiology (2003)
  - Surgical Critical Care (2004)
  - Thoracic Surgery (1988)
  - Vascular Surgery (1988)

References:

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth, "The New York City High School Match," *American Economic Review*, Papers and Proceedings, 95,2, May, 2005, 364-367.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth, "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," working paper, November 2006.

Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez, "The Boston Public School Match," *American Economic Review*, Papers and Proceedings, 95,2, May, 2005, 368-371.

Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez, "Changing the Boston School Choice Mechanism," working paper, January 2006.

Abdulkadiroğlu, Atila and Tayfun Sönmez [1998], " Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems" (joint with Atila), ***Econometrica*** 66: 689-701, May

Abdulkadiroğlu, Atila and Tayfun Sönmez [1999], "House Allocation with Existing Tenants," Journal of Economic Theory, 88, 233-260.

Abdulkadiroğlu, Atila and Tayfun Sönmez [2002], "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729-747.

Abeledo, Hernan G. and Garth Isaak [1991], "A Characterization of Graphs which Assure the Existence of Stable Matchings," *Mathematical Social Sciences*, 22, 93-96.

Abeledo, Hernan G. and Uriel G. Rothblum [1995], "Courtship and Linear Programming," *Linear Algebra and its Applications*, 216: 111-124.

Abraham, D., Blum, A., and Sandholm, T. 2007. Clearing Algorithms for Barter Exchange Markets: Enabling Nationwide Kidney Exchanges. In Proceedings of the *ACM Conference on Electronic Commerce (EC).*

Adachi, Hiroyuki [2000], "On a Characterization of Stable Matchings," *Economic Letters*, 68, 43--49.

Aldershof, Brian and Olivia M. Carducci [1996], ``Stable Matchings with Couples," Discrete Applied Mathematics, (68)1-2 (1996) pp. 203-207.

Alkan, Ahmet [1988], "Nonexistence of Stable Threesome Matchings," *Mathematical Social Sciences*, 16, pp207-209.

Avery, Christopher, Andrew Fairbanks, and Richard Zeckhauser. *The Early Admissions Game: Joining the Elite*. Harvard University Press, 2003.

Avery, Christopher, Christine Jolls, Richard A. Posner, and Alvin E. Roth, "The Market for Federal Judicial Law Clerks", *University of Chicago Law Review*, 68, 3, Summer, 2001, 793-902.

Avery, Christopher, Christine Jolls, Richard A. Posner, and Alvin E. Roth [2007], "The New Market for Federal Judicial Law Clerks", *University of Chicago Law Review*, forthcoming.

Balinski, Michel and Tayfun Sönmez [1999], "A Tale of Two Mechanisms: Student Placement," Journal of Economic Theory, 84, 73-94.

Becker, Gary S. [1981], *A Treatise on the Family*, Cambridge, Harvard University Press.

Blair, Charles [1988], "The Lattice Structure of the Set of Stable Matchings with Multiple Partners," *Mathematics of Operations Research*, 13, pp619-628.

Bulow, Jeremy and Jonathan Levin [2006] "Matching and Price Competition," *American Economic Review*, 96, 3, June, 652-658.

Crawford, Vincent P. "The Flexible-Salary Match: A Proposal to Increase the Salary Flexibility of the National Resident Matching Program," forthcoming, *Journal of Economic Behavior and Organization.*

Crawford, Vincent P. and Elsie Marie Knoer [1981], "Job Matching with Heterogeneous Firms and Workers," **Econometrica**, 49, pp437-450.

Damiano, Ettore, Hao Li, and Wing Suen (2005), "Unravelling of Dynamic Sorting," *Review of Economic Studies* 72 (4), 1057–1076.

Demange, Gabrielle and David Gale [1985] "The Strategy Structure of Two-Sided Matching Markets," *Econometrica*, 53, pp873-888.

Demange, Gabrielle, David Gale, and Marilda Sotomayor [1986] "Multi-Item Auctions," *Journal of Political Economy*, 94, pp863-872.

Dubins, L.E. and D.A. Freedman [1981], "Machiavelli and the Gale- Shapley Algorithm." *American Mathematical Monthly*, 88, pp485-494.

Echenique, Federico and Jorge Oviedo [2004] `Core Many-to-one Matchings by Fixed Point Methods" *Journal of Economic Theory*, 115, 2, April, 358-376.

Echenique, Federico and Jorge Oviedo [2006], "A Theory of Stability in Many-to-many Matching Markets, "*Theoretical Economics,*1, 2, June, 233-273.

Echenique, Federico and M. Bumin Yenmez (2007) , "A solution to matching with preferences over colleagues," *Games and Economic Behavior*, 59, 1, April, 46-71

Ehlers, Lars and Bettina Klaus [2006], "Efficient Priority Rules," *Games and Economic Behavior*, 55, 372-384.

Erdil, Aytek and Haluk Ergin [2007], What's the Matter with Tie-breaking? Improving Efficiency in School Choice, *American Economic Review*, forthcoming.

Erdil, Aytek, and Haluk Ergin [2006] "Two-Sided Matching with Indifferences," working paper.

Ergin, Haluk [2002] "Efficient Resource Allocation on the Basis of Priorities," November, *Econometrica*, 70(6), 2489-2497

Ergin, Haluk and Tayfun Sonmez [2006] "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 90, 215-237.

Fleiner, Tamas [2003], "A Fixed-Point Approach To Stable Matchings and Some Applications," *Mathematics of Operations Research*, 28, 1, February, 103-126

Fréchette, Guillaume, Alvin E. Roth, and M. Utku Ünver, "Unraveling Yields Inefficient Matchings: Evidence from Post-Season College Football Bowls," *Rand Journal of Economics*, forthcoming.

Gale, David and Lloyd Shapley [1962], "College Admissions and the Stability of Marriage,"; *American Mathematical Monthly*, 69, pp9-15.

Gale, David and Marilda Sotomayor [1985] "Ms Machiavelli and the Stable Matching Problem," *American Mathematical Monthly*, 92, pp261-268.

Guillen, Pablo and Onur Kesten [2007], "On-Campus Housing: Theory vs. Experiment," working paper, April, Carnegie Mellon University.

Hatfield, John and Paul Milgrom [2005], "Matching with Contracts," *American Economic Review*, 95(4), September, 913-935.

Immorlica, Nicole and Mohammad Mahdian [2005], "Marriage, Honesty, and Stability," SODA 2005, 53-62.

Irving, Robert W. [1998] "Matching medical students to pairs of hospitals: a new variation on an old theme', in Proceedings of ESA'98, the Sixth Annual European

Symposium on Algorithms, Venice Italy, 1998, Lecture Notes in Computer Science, vol. 1461, Springer-Verlag, 381-392.

Irving RW, Manlove DF, Scott S [2000] The hospitals/residents problem with ties *Lecture Notes in Computer Science* 1851: 259-271.

Kagel, John H. and A.E. Roth, "The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment," *Quarterly Journal of Economics*, February, 2000, 201-235.

Kamecke, Ulrich [1998], "Wage Formation in a Centralized Matching Market," *International Economic Review*, 39, 1, February, 33-53.

Kelso, Alexander S.,Jr. and Vincent P. Crawford [1982], "Job Matching, Coalition Formation, and Gross Substitutes", *Econometrica*, 50, pp 1483-1504.

Kesten, Onur [2004], "Student Placement to Public Schools in US: Two New Solutions," working paper, Carnegie Mellon University.

Kesten, Onur [2006] On Two Competing Mechanisms for Priority Based Allocation Problems, *Journal of Economic Theory* 127, 155-171.

Klaus, Bettina and Flip Klijn (2005): "Stable Matchings and Preferences of Couples," *Journal of Economic Theory*, 121(1), 75-106.

Knuth, Donald E. [1976], *Mariages Stables*, Montreal, Les Presses de l'Universite de Montreal.

Kojima, Fuhito [2007], "Matching and Price Competition: Comment," *American Economic Review*, forthcoming.

Kojima, Fuhito and Parag Pathak [2007], "Incentives and Stability in Large Two-Sided Matching Markets," working paper.

Li, Hao and Sherwin Rosen [1998], "Unraveling in Matching Markets," *American Economic Review*, 88, 3, June, 371-387.

Li, Hao and Wing Suen [2000], "Risk Sharing, Sorting, and Early Contracting," *Journal of Political Economy*, 108, 5, 1058-1091.

Li, Hao and Wing Suen [2004], "Self-Fulfilling Early-Contracting Rush," *International Economic Review*, February, vol. 45, no. 1, pp. 301-324(24)

Martinez, Ruth, Jordi Masso Alejandro Neme, and Jorge Oviedo [2001], On the Lattice Structure of the Set of Stable Matchings for a Many-to-one Model, *Optimization* 50, 439-457.

C. Nicholas McKinney, Muriel Niederle, and Alvin E. Roth, "The collapse of a medical labor clearinghouse (and why such failures are rare)," *American Economic Review*, *95, 3, June, 2005,878-889.*

McVitie, D. G. and L. B. Wilson [1970a], "Stable Marriage Assignments for Unequal Sets," *BIT*, 10, pp295-309.

Milgrom, Paul [2004] *Putting Auction Theory to Work*, Cambridge University Press.

Mullin, F.J. and John M. Stalnaker [1952], "The Matching Plan for Internship Placement: A Report of the First Year's Experience," *Journal of Medical Education*, 27, 193-200.

Niederle, Muriel [forthcoming], "Competitive Wages in a Match with Ordered Contracts," *American Economic Review*.

Niederle, Muriel and Alvin E. Roth, "Relationship Between Wages and Presence of a Match in Medical Fellowships," *JAMA. Journal of the American Medical Association*, vol. 290, No. 9, September 3, 2003, 1153-1154.

Niederle, Muriel and Alvin E. Roth, "Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match," *Journal of Political Economy*, vol. 111, no. 6, December 2003, 1342-1352.

Niederle, Muriel and Alvin E. Roth, "The Gastroenterology Fellowship Market: Should there be a Match?," *American Economic Review, Papers and Proceedings*, 95,2, May, 2005, 372-375.

Niederle, Muriel and Alvin E. Roth, "The Gastroenterology Fellowship Match: How it failed, and why it could succeed once again," *Gastroenterology*, 127, 2 August 2004, 658-666.

Niederle, Muriel and Alvin E. Roth, "The Effects of a Central Clearinghouse on Job placement, Wages, and Hiring Practices: Gastroenterology Fellows as a Case Study," NBER volume, forthcoming.

Niederle, Muriel, Deborah D. Proctor, and Alvin E. Roth, "What will be needed for the new GI fellowship match to succeed?" *Gastroenterology*, January, 2006, 130, 218-224.

Niederle, Muriel and Leeat Yariv [2007], "Matching through Decentralized Markets: Obstacles to Stability," in preparation.

Ostrovsky, Michael, [2007] Stability in Supply Chain Networks, working paper (revised), Stanford University GSB.

Papai, Szilvia [2000] "Strategyproof Assignment by Hierarchical Exchange," *Econometrica*, 68: 1403-1433 .

Pathak, Parag [2007], "Lotteries in Student Assignment," working paper.

Pathak, Parag and Tayfun Sönmez [2007], "Leveling the Playing Field: Sincere and Strategic Players in the Boston Mechanism," working paper.

Public Law 108-218; Pension Funding Equity Act of 2004, SEC. 207., http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=108_cong_public_laws&docid=f:publ218.108.pdf

Pycia, Marek "Many-to-One Matching with Complementarities and Peer Effects," April, 2007, http://www.econ.psu.edu/~pycia/pycia-matching-with-complementarities-April30-07.pdf

Roth, A.E. "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, Vol. 7, 1982a, 617-628.

Roth, A.E. "Incentive Compatibility in a Market with Indivisible Goods", *Economics Letters*, Vol. 9, 1982b, 127-132.

Roth, Alvin E. [1984], "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 92, 991-1016.

Roth, A.E. "The College Admissions Problem is not Equivalent to the Marriage Problem," *Journal of Economic Theory*, 36, 1985, 277-288.

Roth, A.E. "On the Allocation of Residents to Rural Hospitals: A General Property of Two-Sided Matching Markets," *Econometrica*, 54, 1986, 425-427.

Roth, A.E. "New Physicians: A Natural Experiment in Market Organization," Science, 250, 1990, 1524-1528.

Roth, Alvin E. [1991], "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K.," *American Economic Review*, 81, June, 415-440.

Roth, Alvin E. "The Economist as Engineer: Game Theory, Experimental Economics and Computation as Tools of Design Economics," Fisher Schultz lecture, *Econometrica*, 70, 4, July 2002, 1341-1378

Roth, Alvin E. "The origins, history, and design of the resident match," *JAMA. Journal of the American Medical Association*, vol. 289, No. 7, February 19, 2003, 909-912.

Roth, A.E. and E. Peranson, "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89, 4, September, 1999, 748-780.

Roth, Alvin E., Tayfun Sönmez and M. Utku Ünver, "Kidney Exchange," *Quarterly Journal of Economics*, 119, 2, May, 2004, 457-488.

Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver, "A Kidney Exchange Clearinghouse in New England," *American Economic Review,* Papers and Proceedings, 95,2, May, 2005, 376-380.

Roth, Alvin E., Tayfun Sönmez and M. Utku Ünver "Efficient Kidney Exchange: Coincidence of Wants in a Market with Compatibility-Based Preferences," *American Economic Review*, June 2007, forthcoming.

Roth, A.E. and M. Sotomayor "Interior Points in the Core of Two-Sided Matching Markets," *Journal of Economic Theory*, Vol. 45, 1988, 85-101.

Roth, A.E. and Sotomayor, M. "The College Admissions Problem Revisited," *Econometrica*, Vol. 57, 1989, 559-570.

Roth, A.E. and M. Sotomayor *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Econometric Society Monograph Series, Cambridge University Press, 1990.

Roth, A.E. and Vande Vate, J.H. "Random Paths to Stability in Two-Sided Matching," *Econometrica*, 58, 1990, 1475-1480.

Roth, A.E. and X. Xing "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," *American Economic Review*, 84, September, 1994, 992-1044.

Roth, A.E. and X. Xing "Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists," *Journal of Political Economy*, 105, April 1997, 284-329.

Shapley, Lloyd S. and Martin Shubik [1972], "The Assignment Game I: The Core," *International Journal of Game Theory*, 1, pp111-130.

Shapley, Lloyd S. and Herbert Scarf [1974], "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, pp23-28.

Sönmez, Tayfun [1997], "Manipulation via Capacities in Two-Sided Matching Markets," *Journal of Economic Theory*, 77, 1, November, 197-204.

Sönmez, Tayfun [1999], "Can Pre-Arranged Matches be Avoided in Two-Sided Matching Markets?" *Journal of Economic Theory,* 86, 148-56.

Sönmez, Tayfun and M. Utku Ünver [2005], "House Allocation with Existing Tenants: An Equivalence," *Games and Economic Behavior* 52, July, 153-185.

Sotomayor, Marilda [1996], "A Non Constructive Elementary Proof of the Existence of Stable Marriages," Games and Economic Behavior, 13, March, 135-7.

Sotomayor, Marilda [2000], Existence of stable outcomes and the lattice property for a unified matching market, Mathematical Social Sciences, 39, 119-132.

Sotomayor, Marilda [2007], "Connecting the cooperative and competitive structures of the multiple-partners assignment game," *Journal of Economic Theory*, 134, 1, May, 155-174.

Tarski, Alfred [1955], "A Lattice Theoretical Fixpoint Theorem and Its Applications," *Pacific Journal of Mathematics*, 5, 2.

Ünver, M. Utku [2001]"Backward Unraveling over Time: The Evolution of Strategic Behavior in the Entry-Level British Medical Labor Markets" *Journal of Economic Dynamics and Control*, 25, 1039-1080.

Appendix. Some lost and found intellectual history: the Gale-Peranson correspondence

When I began to circulate the material that eventually became Roth (1982a, 1984), David Gale told me that in the 1970's he had corresponded with someone at the medical match, but had not kept any copies of his correspondence. He subsequently told me that Marilda Sotomayor had developed an example like mine showing that the set of stable matchings could be empty when couples were present. However when I asked her about this she told me that this example too had been lost, and in Roth and Sotomayor (1990) we referenced her example as "unpublished."

It turns out that Elliott Peranson had retained copies of correspondence that he shared with me as I prepared this review. It includes, among other things, Gale's 1975 letter of inquiry, Peranson's reply about the version of the deferred acceptance algorithm he had developed in 1972 (recall the history in section 3) which by then already included the "leading member" couples algorithm, and, in a 1983 letter, the lost Sotomayor example. With permission of all concerned (Gale, Peranson, and Sotomayor) I present that correspondence in this Appendix, as a recovered part of the intellectual history of matching and market design, in honor of David Gale's 85th birthday.

National Matching Services Inc.

20 Holly Street, Suite 301
Toronto, Ontario M4S 3B1
Phone: (416) 977-3431
Fax: (416) 977-5020

FAX TRANSMISSION COVER SHEET

Date:	*March 30, 2007*
To:	*Al Roth*
Fax:	*617 495-6537*
Subject:	*Correspondence with David Gale*
Sender:	*Elliott Peranson*

YOU SHOULD RECEIVE 12 PAGE(S), INCLUDING THIS COVER SHEET.
IF YOU DO NOT RECEIVE ALL THE PAGES, PLEASE CALL.

Al,

This is all the correspondence I had with David Gale that I could find. I think there was more, but it has disappeared over time.

I have included a later correspondence from 1983 about couples, because I thought you might find it interesting, for two reasons - it indicates what appears to be the first time he considered matches that had no stable solution, and it refers to Marilda Sotomayor.

Elliott

CENTER FOR ADVANCED STUDY IN THE BEHAVIORAL SCIENCES

202 Junipero Serra Boulevard • Stanford, California 94305 Telephone (415) 321-2052

6 November 1975

Ms. Alice Skarvinsky
1603 Orrington Avenue
Evanston, Illinois 60201

Dear Ms Skarvinsky:

Following our phone conversation of yesterday, I am enclosing a copy of the paper I mentioned to you which deals with the matching problem. Of course it would be interesting to me to know if your procedure is the same as the one described in our article which we show is the only possible one which is "reasonable" (in the sense of stability) and which favors the students over the hospitals. However, I feel it is rather an imposition to ask people to spend time reading the paper. Perhaps, though, I could ask one very specific question which I do not think would take up too much of anyone's time. Attached are two examples of a hypothetical situation involving four students and four hospitals in which each hospital has a quota of one. I think the notation is clear. Each entry in the table consists of a pair of numbers. The first entry in row i, column j gives the ranking of hospital j by student i and the second number gives the ranking of student i by hospital j. Thus in table one, student 3 ranks hospital 2 in first place while hospital 2 ranks student 3 in 4th place, etc. My question is simply, how would your algorithm assign students to hospitals in each of the two examples? For such small examples a computer is of course not needed and I expect you or one of your technical people could provide this information with only a few minutes of hand calculation.

I'm sorry to trouble you with this request, but every time I talk on the subject someone always mentiones the National Matching Program and I feel it's time I found out what the relationship is in the instance between "theory and practice."

(My current address is given on the above letterhead. I am on leave from my regular job in the Mathematics Department at the University of California, Berkeley.)

Sincerely yours,

David Gale

David Gale

DG:mt

TABLE I

	H_1	H_2	H_3	H_4
S_1	1,3	2,3	3,2	4,3
S_2	1,4	4,1	3,3	2,2
S_3	2,2	1,4	3,4	4,1
S_4	4,1	2,2	3,1	1,4

TABLE II

	H_1	H_2	H_3	H_4
S_1	1,3	2,2	3,~~2~~	4,3
S_2	1,4	2,3	3,2	4,4
S_3	3,1	1,4	2,3	4,2
S_4	2,2	3,1	1,4	4,1

CENTNER PERANSON AND ASSOCIATES

122 Cumberland St.,
Toronto, Canada M5R 1A6
(416) 961-3660

December 8, 1975

Dr. David Gale
Center for Advanced Study in the Behavioural Sciences
202 Junipero Serra Boulevard
Stanford, California 94305

Dear Dr. Gale:

As I am the consultant to NIRMP responsible for the technical operation of the matching program, Ms. Skarzynski has referred your recent letter and request to me. I apologize for not responding earlier but we are currently in the middle of a hectic period in NIRMP's schedule of activities.

Enclosed please find the two tables you sent to NIRMP. The results of the matching algorithm used by NIRMP are indicated by the red circles on the tables. For example, in Table I student S_1 is matched to Hospital H_3, etc.

You will note that the results indicated here match those that would be obtained from the "deferred - acceptance" algorithm outlined in your paper. However, I might point out that the NIRMP algorithm in fact uses the inverse procedure and produces the unique "college optimal" assignment rather than the "student optimal" assignment. This procedure more closely parallels the actual admissions process where a matching algorithm is not used. In this case students apply to all hospitals they would consider (not just their first choice), each hospital then selects the most desirable students, up to its quota limit, from amongst all applicants, then the "bid-for" students reject all but the most desirable offer, and so on. The use of the "college optimal" assignment algorithm was approved by the students when NIRMP first began its operation twenty five years ago.

Having answered your request I would now like to pose a problem to you which is, to me, a much more academically challenging problem than that of devising a basic matching algorithm. This problem stems from the commitment NIRMP makes to its married or engaged applicants to match two students (married or engaged) to hospitals in the same community, while maintaining the overall optimality and integrity of the results of the match. Married couples are permitted to apply separately to any hospital programs of their choice, they are permitted to prepare their rank order

.../2..

lists as individuals (e.g. the two students can assign completely different ranks to the same hospital or two hospitals in the same community), hospitals may rank the two students as individuals (e.g. assign one a high rank and one a low rank), and each couple can even define for itself which hospitals it considers to be in the same community. A copy of the instructions sent to couples outlining the additional data we receive from married or engaged couples is enclosed with this letter.

In order to handle the situation of married or engaged applicants as described above and still produce an optimal match for all applicants, singles as well as couples, we recently developed a two stage computerized matching procedure. In this first stage a modification of the original matching algorithm is combined with a set of decision rules for couples to produce a result which matches all singles and about 98% of all couples in an optimal fashion. The second stage then analyzes the specific non-optimal couples matches to identify possible improvements, adjusts the results accordingly and then iterates to a final optimal solution. While this procedure does produce an optimal match result, it must resort to a somewhat "brute force", inelegant approach; it would be interesting from an academic viewpoint, and perhaps even more efficient in practice, to have a single, programmable algorithm which produced an overall optimal match directly in the first stage. If you wish to consider this additional wrinkle in the match procedure and can derive an efficient algorithm to solve this problem I would be interested in hearing from you.

In any case, I hope this letter and the accompanying tables has answered any questions you have concerning the basic matching algorithm for single applicants currently used by NIRMP.

Yours truly,

Elliott Peranson

Elliott Peranson

Enclosures

TABLE I

	H_1	H_2	H_3	H_4
S_1	1,3	2,3	(3,2)	4,3
S_2	1,4	4,1	3,3	(2,2)
S_3	(2,2)	1,4	3,4	4,1
S_4	4,1	(2,2)	3,1	1,4

TABLE II

	H_1	H_2	H_3	H_4
S_1	1,3	2,2	(3,[redacted])	4,3
S_2	1,4	2,3	3,2	(4,4)
S_3	(3,1)	1,4	2,3	4,2
S_4	2,2	(3,1)	1,4	4,1

CENTER FOR ADVANCED STUDY IN THE BEHAVIORAL SCIENCES

202 Junipero Serra Boulevard • Stanford, California 94305

Telephone (415) 321-2052

5 January 1976

Mr. Elliott Peranson
Centaer Peranson and Associates
122 Cumberland Street
Toronto, Canada M5R 1A6

Dear Mr. Peranson:

Many thanks for your letter and reply to my questions regarding your matching program. The relation between your program and ours is very interesting.

Regarding the problem of matching couples it seems much more complicated. It's not clear to me in this case what is meant by an optimal matching. In order to think about the problem I would have to have a precise definition of "optimality." Would an optimal matching for this case be unique as it is in the single applicant case? If you could supply me with this definition I might be able to give the question some further thought.

Sincerely yours,

David Gale

DG:mt

UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA

DEPARTMENT OF MATHEMATICS

BERKELEY, CALIFORNIA 94720

August 10, 1983

'83 AUG 15 11

Mr. Elliot Parenson
National Resident Matching Program
One American Plaza
Suite 807
Evanston, Illinois 60201

Dear Mr. Parenson:

I appreciated our informative telephone conversation the other day. I mentioned your problem of matching couples to a young woman, Marilda Sotomay a visiting mathematician from Brazil. She came up with what strikes me as an interesting example in which there is no "stable" matching in the sense of our original paper. That is, no matter how you assign students to hospitals it will always be the case that there will be some hospital H and student α belonging to the couple (α,β) such that H prefers α to the student assigned to it and the couple (α,β) prefers (H,H_β) to (H_α,H_β) where H_α and H_β are the hospitals to which α and β have been assigned. The example is of course highly contrived but it shows the possibility of instability. The enclosed pages show on page 0 the rank orderings of the four hospitals and the two couples. Each hospital (or program, if you prefer) has a quota of 1. The following pages give all 24 possible matchings and prove their instability. Thus, on the top line of page 1 m_1 goes to H_1, w_1 goes to H_2, m_2 goes to H_3, w_2 goes to H_4. The line connecting H_1 to w_2 shows the instability. It means that H_2 prefers w_2 to w_1, and as shown on the right $\langle m_2,w_2\rangle$ prefer (H_3,H_2) to (H_3,H_4) and so on.

I would be curious to know what your matching algorithm does in this example. Does it simply run forever without terminating or does it pick out one of the unstable matchings?

I can see how introducing couples does substantially complicate the whole problem. I don't even know what one would mean by an optimal matching, say in the case of the above example.

I will be interested in your reaction to this.

Sincerely yours,

David Gale

David Gale

DG:gy
cc: Phyllis Wesland

Couples: $\{\langle m_1, w_1\rangle, \langle m_2, w_2\rangle\}$

Hospitals: $\{H_1, H_2, H_3, H_4\}$.

H_1: m_1 w_2 w_1 m_2

H_2: m_1 w_2 w_1 m_2

H_3: m_2 w_1 m_1 w_2

H_4: m_2 w_2 m_1 w_1

$\langle m_1, w_1\rangle$: (H_2, H_1), (H_1, H_3), (H_4, H_1), (H_4, H_3), (H_1, H_2), (H_2, H_4)

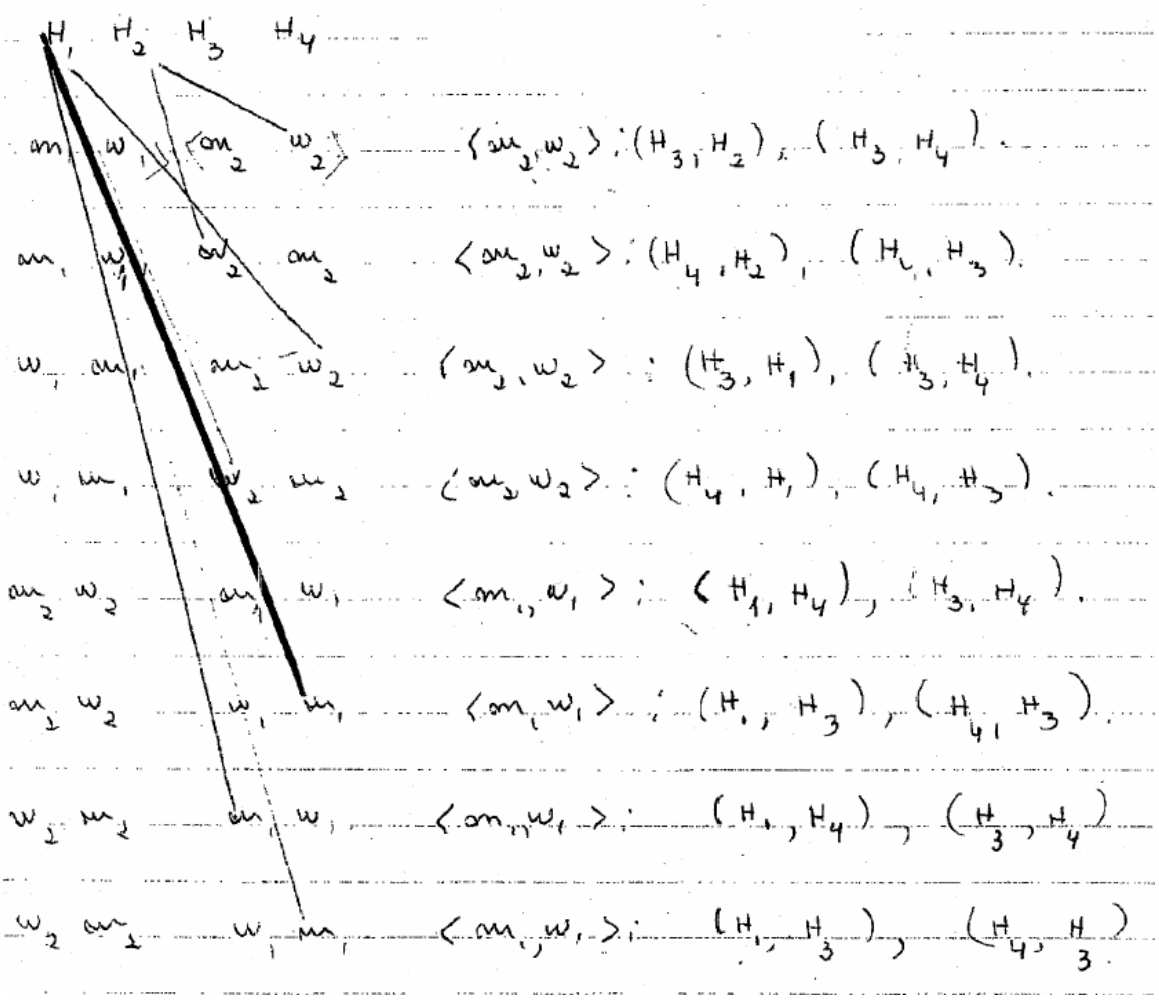
(H_1, H_4), (H_3, H_2), (H_3, H_4), (H_2, H_3), (H_4, H_2)

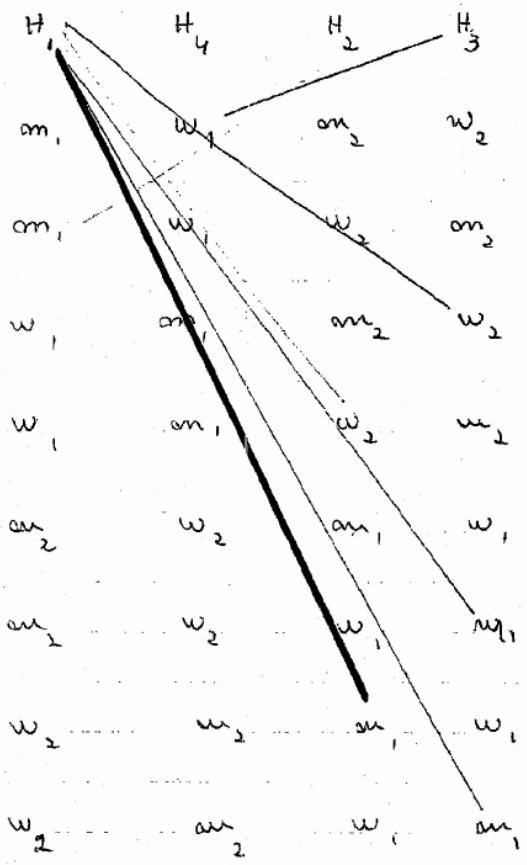
$\langle m_2, w_2\rangle$: (H_4, H_1), (H_3, H_1), (H_2, H_1), (H_3, H_2), (H_3, H_4), (H_2, H_4)

(H_2, H_3), (H_4, H_2), (H_4, H_3), (H_1, H_4), (H_1, H_3)

NATIONAL RESIDENT MATCHING PROGRAM

'83 AUG 15 11:1


H1 H2 H3 H4
m1 w1 m2 w2 ⟨m2, w2⟩ : (H3, H2), (H3, H4)
m1 w1 w2 m2 ⟨m2, w2⟩ : (H4, H2), (H1, H3)
w1 m1 m2 w2 ⟨m2, w2⟩ : (H3, H1), (H3, H4)
w1 m1 w2 m2 ⟨m2, w2⟩ : (H4, H1), (H4, H3)
m2 w2 m1 w1 ⟨m1, w1⟩ : (H1, H4), (H3, H4)
m2 w2 w1 m1 ⟨m1, w1⟩ : (H1, H3), (H4, H3)
w2 m2 m1 w1 ⟨m1, w1⟩ : (H1, H4), (H3, H4)
w2 m2 w1 m1 ⟨m1, w1⟩ : (H1, H3), (H4, H3)



$\langle m_1, w_1 \rangle : (H_1, H_3), (H_1, H_4)$.

$\langle m_1, w_1 \rangle : (H_2, H_4), (H_1, H_4)$.

$\langle m_2, w_2 \rangle : (H_2, H_1), (H_2, H_3)$

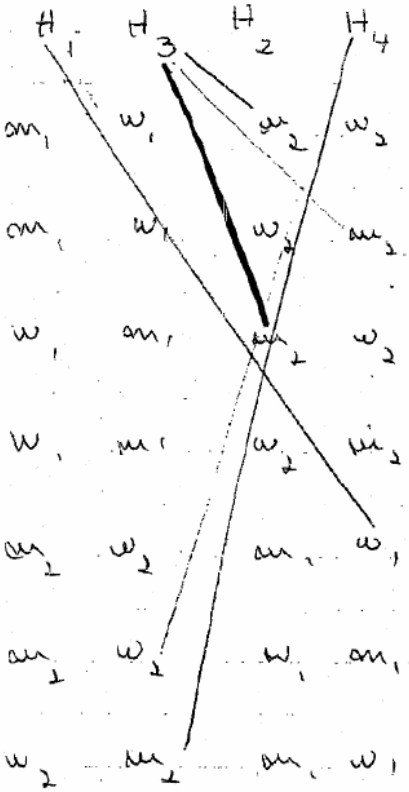
$\langle m_2, w_2 \rangle : (H_3, H_1), (H_3, H_2)$

$\langle m_1, w_1 \rangle : (H_1, H_3), (H_2, H_3)$

$\langle m_1, w_1 \rangle : (H_1, H_2), (H_3, H_2)$

$\langle m_1, w_1 \rangle : (H_1, H_3), (H_2, H_3)$

$\langle m_1, w_1 \rangle : (H_1, H_2), (H_3, H_2)$



m_1 w_1 m_2 w_2 $\langle m_2, w_2 \rangle : (H_3, H_4), (H_2, H_4)$.

m_1 w_1 w_2 m_2 $\langle m_2, w_2 \rangle : (H_3, H_2), (H_4, H_2)$.

w_1 m_1 m_2 w_2 $\langle m_2 w_2 \rangle : (H_3, H_4), (H_2, H_4)$.

w_1 m_1 w_2 m_2 $\langle m_2, w_2 \rangle ; (H_3, H_2), (H_4, H_2)$.

m_2 w_2 m_1 w_1 $\langle m_1, w_1 \rangle : (H_2, H_1), (H_2, H_4)$.

m_2 w_2 w_1 m_1 $\langle m_2, w_2 \rangle : (H_1, H_4), (H_1, H_3)$.

w_2 m_2 m_1 w_1 $\langle m_2, w_2 \rangle : (H_4, H_1), (H_3, H_1)$.

w_2 m_2 w_1 m_1 $\langle m_2, w_2 \rangle : (H_4, H_1), (H_3, H_1)$.